

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: November 27, 2002, 02:06:46 ; Search time 2611 Seconds
(without alignments)
12104.805 Million cell updates/sec

Title: US-09-374-967-1
Perfect score: 1086
Sequence: 1 atgaagccctcattctgt.....agcctgagatcgtcatgtga 1086

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :		GenEmbl.*
1:	gb.ba.*	
2:	gb.htg.*	
3:	gb.in.*	
4:	gb.om.*	
5:	gb.ov.*	
6:	gb.pat.*	
7:	gb.ph.*	
8:	gb.pl.*	
9:	gb.pr.*	
10:	gb.ro.*	
11:	gb.sts.*	
12:	gb.sy.*	
13:	gb.un.*	
14:	gb.vi.*	
15:	em.ba.*	
16:	em.fun.*	
17:	em.hum.*	
18:	em.in.*	
19:	em.mu.*	
20:	em.om.*	
21:	em.or.*	
22:	em.ov.*	
23:	em.pat.*	
24:	em.ph.*	
25:	em.pl.*	
26:	em.ro.*	
27:	em.sts.*	
28:	em.un.*	
29:	em.vi.*	
30:	em.htg_hum.*	
31:	em.htg_inv.*	
32:	em.htg_other.*	
33:	em.htg_mus.*	
34:	em.htg_pln.*	
35:	em.htg_rod.*	
36:	em.htg_mam.*	
37:	em.htg_vrt.*	
38:	em.sy.*	
39:	em.htgo_hum.*	
40:	em.htgo_mus.*	
41:	em.htgo_other.*	

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	692.4	63.8	1440	8	AF428297	Arabidops
2	690.8	63.6	1086	8	AF133643	Arabidops
3	690.8	63.6	1450	8	AF361812	Arabidops
4	690.8	63.6	1477	8	AF108660	Arabidops
5	690.8	63.6	1480	8	AY087698	Arabidops
6	690.8	63.6	1482	8	AY057541	Arabidops
7	690.8	63.6	1490	8	AF076484	Arabidops
8	653	60.1	142885	8	AP004127	Oryza sat
9	653	60.1	149145	8	AP003251	Oryza sat
10	651.4	60.0	142024	2	AP004705	Oryza sat
11	650.8	59.9	1478	6	AY7607	Sequence 1
12	650.8	59.9	1478	8	AF022716	Solanum t
13	638	58.7	1471	8	AB066279	Nicotiana
14	486	44.8	92624	8	AC003000	Arabidops
15	461.2	42.5	103240	8	ATT22E16	Arabidops
16	461.2	42.5	103785	8	ATF1116	Arabidops
17	451	41.5	9423	8	ATH275979	Arabidops
18	417.2	38.4	1656	3	AY071411	Drosophil
19	417.2	38.4	1715	3	AY061013	Drosophil
20	399.6	36.8	1498	9	AF135421	Homo sapi
21	399.6	36.8	1555	9	BC001141	Homo sapi
22	399.6	36.8	1555	9	BC008033	Homo sapi
23	399.6	36.8	6077	9	AB058754	Homo sapi
24	372.4	34.3	1619	8	AF234177	Pichia an
25	366	33.7	1785	8	HJ089991	Hypocrea je
26	365.8	33.7	3267	8	SCU19608	Saccharomyc
27	365.8	33.7	4446	8	SCYDL055C	S.cerevisia
28	364.2	33.5	2745	8	SCU24437	Saccharomyc
29	358	33.0	1092	6	AX250395	Sequence
30	358	33.0	15252	8	SPCC1906	Sequence
31	357.4	32.9	3220	9	AK024319	S.pombe c
32	356.2	32.8	2179	8	AB020570	Candida g
33	354.8	32.7	1275	8	D89128	Schizosacch
34	350.8	32.3	134784	8	F17123	Arabidops
35	350.8	32.3	194862	8	ATCHR173	Arabidops
36	312.4	28.8	1583	8	AF030299	Candida a
37	310.8	28.6	1089	6	AX488923	Sequence
38	306	28.2	1593	8	AF030300	Candida a
39	304.4	28.0	1665	8	AB020596	Candida a
40	282.2	26.0	144483	2	AC091701	Trypanoso
41	269.4	24.8	40519	3	LMFL2581	Leishmani
42	261.4	24.1	44686	3	AF043695	Caenorhab
43	257	23.7	2199	3	LME292039	Leishmani
44	240	22.1	267509	3	CNS07EGH	chromosom
45	215.6	19.9	45701	2	AC014674	Drosophil

ALIGNMENTS

RESULT 1	AF428297	1440 bp	mrna	linear	PLN 18-OCT-2001
LOCUS	Arabidopsis thaliana	At2g39770/T517.7	mrna	complete cds.	
DEFINITION	Arabidopsis thaliana	At2g39770/T517.7	mrna	complete cds.	
ACCESSION	AF428297	GI:16226304			
VERSION	AF428297.1	FLI_CDNA			
KEYWORDS	FLI_CDNA				
SOURCE	Arabidopsis thaliana				
ORGANISM	Arabidopsis thaliana				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis				
AUTHORS	1 (bases 1 to 1440)				
	Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J., Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,				

RESULT 2	AY133643	1086 bp	linear	PLN 07-AUG-2002	VEKPKLYGVGNKINAGIYLLNPVLDKIELRPTSEIKETFPKIAAAGLYAWLPQFM
LOCUS	Arabidopsis thaliana At2g39770/T517.7 mRNA, complete cds.				DIGQPRDYITGLRLYSLRKKSPAKLTSGPHVNVLTDEATIGSGCLIGPDVAIG
DEFINITION	Arabidopsis thaliana At2g39770/T517.7 mRNA, complete cds.				PGCIVESGVRLSCTVMRGVRIKKHACISSIIIGWHSTVGQWARIENWTILGEDVHVS
ACCESSION	AY133643				DEIYNGGVLPKHKEIKSNILKPEIYM"
KEYWORDS	FLI_CDNA.				
SOURCE	thale cress.				
ORGANISM	Arabidopsis thaliana				
REFERENCE	1 (bases 1 to 1086)				
AUTHORS	Chen, R., Chen, H., Kim, C.J., Shinn, P., Banh, J., Bowser, L., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.C., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A., and Ecker, J.R.				
TITLE	Arabidopsis ORF clones				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 1086)				
AUTHORS	Chen, R., Chen, H., Kim, C.J., Shinn, P., Banh, J., Bowser, L., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.C., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A., and Ecker, J.R.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-JUL-2002) Salk Institute Genomic Analysis Laboratory (SIGnAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA				
COMMENT	RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-Length cDNA') : Seki, M., Narusaka, M., Ishida, J., Sacou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.				
FEATURES	The Salk, Stanford, PGSC (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAFL cDNAs: Cheuk, R., Chen, H., Kim, C.J., Shinn, P., Banh, J., Bowser, L., Chan, M.M., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.C., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Davis, R.W., Theologis, A., and Ecker, J.R.				
source	Location/Qualifiers				
	1..1086				
	/organism="Arabidopsis thaliana"				
	/db_xref="taxon:3702"				
	/chromosome="2"				
	/clone="U14534"				
	/note="This clone is in PUNI 51"				
	ecotype: Columbia"				
	1..1086				
	/note="GDP-mannose pyrophosphorylase"				
	/codon_start=1				
	/product="At2g39770/T517.7"				
	/protein_id="RAM91473.1"				
	/db_xref="GI:22137256"				
	/translation="MKALILVGGFGTRLRPLTLSPFKPLVDFANKPMLHQIEALKAVGVDFVLAINQPVLMFLKDFETKLEIKITCSQETPLGTAGPLALARKLLDGSGEPPFVLNSDVISEVPLKEMLEFHKSHGGEASIMVTKVDEPSKYGVVMEESTGRVEKF				
CDS					
	1..1086				
	/note="GDP-mannose pyrophosphorylase"				
	/codon_start=1				
	/product="At2g39770/T517.7"				
	/protein_id="RAM91473.1"				
	/db_xref="GI:22137256"				
	/translation="MKALILVGGFGTRLRPLTLSPFKPLVDFANKPMLHQIEALKAVGVDFVLAINQPVLMFLKDFETKLEIKITCSQETPLGTAGPLALARKLLDGSGEPPFVLNSDVISEVPLKEMLEFHKSHGGEASIMVTKVDEPSKYGVVMEESTGRVEKF				

BASE COUNT	299 a	220 c	285 g	282 t	
ORIGIN					
Query Match	63.6%	Score 690.8;	DB 8;	Length 1086;	
Best Local Similarity	77.3%	Pred. No. 3.8e-204;			
Matches	839;	Conservative	0;	Mismatches 247;	Indels 0;
Gaps	0;				
QY	1	ATGAAGCCCTCATCTCTGCGGGGGTTCGGAACCGCCTTCGGCCCTTGACTCTGAGC	60		
Db	1	ATGAAGCCCTCATCTCTGCGGGGGTTCGGAACCGCCTTCGGCCCTTGACTCTGAGC	60		
QY	61	TTCCCGAAGCCCTCTGCGGGGGTTCGGAACCGCCTTCGGCCCTTGACTCTGAGC	120		
Db	61	TTCCCGAAGCCCTCTGCGGGGGTTCGGAACCGCCTTCGGCCCTTGACTCTGAGC	120		
QY	121	TTGAAGAAGCTTGGGTCACAGAGGTGTTTGGCTATCAACTATCGCCAGAGGTAATG	180		
Db	121	CTTAAGCCAGTTGGAGTTGATGAAGTGGTTTGGCCATCAATTATCAGCCAGAGGTGATG	180		
QY	181	ATTAATTTCTTGAGGACTTTTGAGGATAAGCTTGGCATCAACATTCATGTCCTCAAGAG	240		
Db	181	CTGAACCTCTTGAGGACTTTTGAGGATAAGCTTGGCATCAACATTCATGTCCTCAAGAG	240		
QY	241	ACTGAGCCCTTAGAAGCCGCTGGCCCTTCTGCTTAGCAAGGACCAAGCTTGGCGATGGA	300		
Db	241	ACCGAGCCACTAGGTACCGCTGGCTCTGAGGAGAGACAAATTCCTTGATGGA	300		
QY	301	TCGCGCCAGCCATCTCTTGTCTCAACAGTGTATGATCAAGCAATACCCATTTGCTGAA	360		
Db	301	TCGCGAGGCGCTCTTGTCTCAACAGTGTATGATCAAGCAATACCCATTTGCTGAA	360		
QY	361	CTCATCAAAATTCACAAAGTGTATGTTGGTGGAGCAACAATATGTCATTAAGAGT	420		
Db	361	ATGCTTGAGTTTCAAAATCTACGGTGGGGAAGCTCCATAATGGTAACAAAGTGGAT	420		
QY	421	GAACCATCAAAATACGGTGTGTTGTTATGGAGGAGCAACTGGCAGGTTGGAAGGTTT	480		
Db	421	GAACCGTCGAAATATGAGTGGTGTGTTATGGAAGAAAGCACTGGAAGAGTGGAGAAGT	480		
QY	481	GTTGAGAAGCCAAATATTTGGGTAAACAAGATCAATGCTGGATTACTTACTGAAC	540		
Db	481	GTGGAAGCCAAATCTGTTAGGTACAAAGATCAACGCTGGATTACTTCTGAAAC	540		
QY	541	CCATCTCTCTTGACCCGATTTAGCTGAGGCAACATCAATTTGAGAAAGAGTCTTCCT	600		
Db	541	CCATCTCTCTTGATTAAGATTGAGCTAAGACCGACTTCAATCGAAAAAGAGACTTTCCT	600		
QY	601	CAATTCAGCTGATCAACAGCTCTATGCAATGCTCTCCAGTTTTCAGGTGATGATGTT	660		
Db	601	AAGATTGACGACGCGAAGGCTCTATGCTATGTTGTTACCCAGGTTTGGATGACAT	660		
QY	661	GGTCAGCTAGGACTACATTTACTGCTTGGCTTGTATCTAGACTCGATTAGGAAGAA	720		
Db	661	GGCAACCCGCTGACTACATAACGGGTTTGAGACTCTACTTAGACTCCCTTAGGAAGAA	720		
QY	721	TCAGCTGCCAAGCTAGCTACTGGAGCACATGTTTGGCAATGTCTGGTGCATGAGAGC	780		
Db	721	TCCTCTGCCAATTAACAGTGGGCCACACATAGTTGGGAATGTTCTTTGACGAAACC	780		
QY	781	GCCAAAGATTGGGAAGTGTCTGATTGGCTCTGATGTCGCCATGGACCTGGGTGTT	840		
Db	781	GCTACAATTTGGGAAGGATGTTTATTTGGACACAGCTTGCATTTGGTCCAGGCTGCA	840		
QY	841	GTGAGGACCGCTGAGGCTTTCCCGCTGCACTCTCATGCGCGCTCGCTATCAAGAG	900		
Db	841	GTTGAGTCAGAGTCAGACTCTCCCGATGACGCTCATGCGTGGAGTCCGATCAAGAG	900		
QY	901	CATGCTTGCACTCTCAAAACAGCATTTATCGGTGGGCACTCAACTGTTGGTCAATGGCAGG	960		


```
Db 879 TCTCTGCCAAATTAACAGTGGGCCACACATAGTTGGGAATGTTCTTTGTGACAAACC 938
QY 781 GCCAAGATTGGAGAAGGTGTCTGATTTGGTCTGATGTCGCATTTGGACCTGGGTGTGTT 840
Db 939 GCTACAATTGGGAAGGATGTTTGAATGGACACAGCTGGCCATTTGGTCCAGGCTGCATT 998
QY 841 GTGAGGACGGCGTAGAGCTTTCCGCTGTCACACTGTCATGCGCGGGCGTGCATCAAGAAG 900
Db 999 GTTGAGTCAGGAGTCAGACTCTCCGATGACCGGTCAATGCGTGGAGTCCGCATCAAGAAG 1058
QY 901 CATGCTTGCATCTCAACAGCATTTATCGGCTGGCGACTCAACTGTGGTCAATGGCCACGG 960
Db 1059 CATCGGTGTAFTCTCGAGCAGTATCATCGGTGGCCACTCAACGGTGGTCAATGGGCCAGG 1118
QY 961 ATAGAGAAATGACTATCTCTGGGGAGGATGTTTATGCTGTGTGATGAGTGTACAGCAAT 1020
Db 1119 ATCAGAACATGACATCTCGGTGAGGATGTTTATGCTGTGAGGATGAGATCTATAGCAAT 1178
QY 1021 GCGCGTGTGTTCTCCACATAAAGAGATCAAGTCAAGCATCTCGAAGCCGTGAGATCGTC 1080
Db 1179 GGAGGAGTGTGTTTGCCACACAAAGGAGATCAAAATCAAAATCTTTGAAGCCAGAGATAGTG 1238
QY 1081 ATGTGA 1086
Db 1239 ATGTGA 1244

RESULT 4
AF108660
LOCUS AF108660 1477 bp mRNA linear PLN 13-JAN-1999
DEFINITION Arabidopsis thaliana Cytl protein (cytl) mRNA, complete cds.
ACCESSION AF108660
VERSION AF108660.1 GI:4151924
KEYWORDS
SOURCE Arabidopsis thaliana.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 1477)
Lukowitz,W. and Somerville,C.
Unpublished
Positional cloning of the Arabidopsis cytl gene
2 (bases 1 to 1477)
Lukowitz,W. and Somerville,C.
Direct Submission
Submitted (21-NOV-1998) Department of Plant Biology, Carnegie
Institution of Washington, 260 Panama Street, Stanford, CA
94305-1297, USA
FEATURES
source
1..1477
/organism="Arabidopsis thaliana"
/cultivar="Columbia"
/db_xref="taxon:3702"
/chromosome="2"
/clone="EST T21688"
1..1477
/gene="cytl"
151..1236
/gene="cytl"
/notes="putative mannose-1-phosphate guanylttransferase"
/codon_start=1
/product="Cytl protein"
/protein_id="A004627.1"
/db_xref="GI:4151925"
translation="MKALILVGGFTRLRPLTSLFPPKPLVDFANKPMLHIOIEALKAV
GVDEVLAINVQPEVLMFLKDETKLEIKITCSQETEPGLGTAGPLARLDKLDGSG
EPFVLNSDVISYPLKEMLEFHKSHGSEASIMVTKVDEFSKYGVVMEESTGRVEKF
VERPKLYGNKLINAGLYLLNPNLSDIKELPTSTIEKTFPKIAAOGLYAMVLPGFWM
DIGQPRDITGLRLYDLSLRKSPAKLTSGPHIVGNVLVDETATIGEGCLIGPDVAIG
PGCIVSGVRLSCTVMRGVRIKKHACISSIIIGWHSTVGQWARIENWTLIGEDVHS
DEIYSGGVLPKHEIKSNILKPEIYM"
```

```
BASE COUNT 397 a 291 c 344 g 445 t
ORIGIN
Query Match 63.6%; Score 690.8; DB 8; Length 1477;
Best Local Similarity 77.3%; Pred. No. 4e-204;
Matches 839; Conservative 0; Mismatches 247; Indels 0; Gaps 0;
QY 1 ATGAAGCCCTCATTTCTGCGGGGGTTTCGGAAACCGCCTTCGGCCTTTGACTCTGAGC 60
Db 151 ATGAAGGCACCTCATTTCTGTTGGAGGCTTCGGCACTCGCTTGAGACCATTTGACTCTCAGT 210
QY 61 TTCCCGAAACCCCTCTGTGATTTTGCAACAAGCCCATGATTTGTCACACAGATCGAAGCT 120
Db 211 TTCCCAAGCCCTCTGTTGATTTGCTAATAAACCCCATGATCTTTCATCAGATAGAGCT 270
QY 121 TTGAAGAAGATTTGGGGTTCACAGAGGTGGTTTGGCTTATCAACTATCGCCCAAGAGTAATG 180
Db 271 CTTAAGGCAGTTGGAGTTGATGAAGTGGTTTGGCCCATCAATTTATCAGCCAGAGGTGATG 330
QY 181 ATTAATTTCTGAAGGACTTTGAGGATAAGCTTTGGCATCACAAATACATGCTCCCAAGAG 240
Db 331 CTGAACCTTCTGAAGGACTTTGAGACCAAGCTGGAATCAAAATCACTTCTCACAAGAG 390
QY 241 ACTGAGCCCTTAGGAACCGCTGCCCTCTTGTCTTAGCAAGGACAAAGCTTGGCGATGGA 300
Db 391 ACCGAGCCACTAGGTACCGCTGGCTCTTGGCTCTTAGCGAGACAAATTTGCTTGTATGGA 450
QY 301 TCTGCCACCCATTTCTTGTCTCAACAGTGTATAGCGAATACCAATTTGCTGAA 360
Db 451 TCTGGAGACCCCTCTTGTCTTAACAGTGTATGATGATGAGTACCCCTCTTAAAGAA 510
QY 361 CTCATCAAAATTTCAACAAGTGTATGTTGGTGGAGCAACAATTTATGGTCACTAAGGTGGAT 420
Db 511 ATGCTTGAGCTTTCACAAATCTCACGCTGGGAGCCCTCCATAATGTTAACAAGGTGGAT 570
QY 421 GAACCATCAAAATACGCTGTTGTGTTATGGAGGAGGCAACTGGCAGGTGGAAAGATTTT 480
Db 571 GAACCGCTCAAAATATGAGGTGGTGTATGGAAGAAAGCACTGGAAGAGTGGAAAGATTTT 630
QY 481 GTTGAGAACCCAAAATATTTGTGGGTAAACAGATCAATGCTGGGATTTACTTACTGAAC 540
Db 631 GTGGAAGCCAAAATGTATGTAGGTAAACAGATCAACGCTGGGATTTATCTCTGAAAC 690
QY 541 CCATCTGTCTTGTACCGCATTTGAGCTGAGGCCCAACATCAATTTGAGAAAGAGGTCTCCCT 600
Db 691 CCATCTGTCTTGTATAAGATTGAGCTAAGCCGACATCAATCGAAAGAGAGACTTTCCCT 750
QY 601 CAAATTTGACGTGATCAACAGCTCTATGCAATGGTCCCTTCCAGGTTTTTGGATGGATGTT 660
Db 751 AAGATTTGACGACGCGCAAGGGCTCTATGCTATGGTGTCTACCGAGTTTTTGGATGGACAT 810
QY 661 GGTACGCCCTTAGGACTACATTTACTGGCTTGGCTTGTGCTTTATCTAGACTCGATTAGGAAGA 720
Db 811 GGGCAACCCCGTACATACATAAGCGGTTTGGAGACTCTACTTACACTCCCTTAGGAAGAAA 870
QY 721 TCAGCTGCCAAGCTAGCTACTGGAGCACATGTTTGTGGCAATGCTGTGGTGCATGAGAGC 780
Db 871 TCTCTGCCCAATTAACAGTGGGCCACACATAGTTGGGAATGTTCTTGTGTAGCAAGACC 930
QY 781 GCCAAGATTTGGAGAAGGTTGTCTGATTGGTTCCTGATGTGCCCATTTGGACCTGGGTGTGTT 840
Db 931 GCTACAATTTGGGAAGGATGTTTGAATTGGACACAGAGCTTGGCCATTTGTCTCCAGGCTGCATT 990
QY 841 GTGGAGGACGGGTGAGGCTTTCCCGCTGCAGCTGTATGCGCGGGGTGCGCTATCAAGAAG 900
Db 991 GTTGAGTCAGGAGTCAGACTCTCCCGATGCACGCTCATGCTGGAGTGGAGTCCGCATCAAGAAG 1050
QY 901 CATGCTTGCATCTCAACAGCATTTATCGGCTGGCATCAACTGTTGGTCAATGGCACGG 960
Db 1051 CATCGGTGTATCTCGAGCAGTATCATCTGGGTGGCAGTCAACGGTGGTCAATGGGCCAGG 1110
QY 961 ATAGAGAAATGACTATCTCTGGGGAGGATGTTTCATGTGTGTGATGAGGTGTACAGCAAT 1020
```


QY	721	TCAGCTGCCAAGCTAGCTACTGGAGCACATGTTGTTGGCAATGCTGCTGATGAGACC	780
Db	873	TCTCTCTGCAATTAACACAGTGGGCCACACATAGTAGTGGAAATGTTCTTTGACGAAACC	932
QY	781	GCCAAAGATTGGAGAGGTTGCTGATTGCTGATGTCGCTGATGCGCCATTTGGACCTGGGTGTGT	840
Db	933	GCTACAATTGGGAAGGATGTTGATTGGACAGACAGTTCGCTGTTGGTCCAGGCTGCATT	992
QY	841	GTGGAGGACGCGGTGAGCGTTTCCGCGTGCACGTGTCATCGCGCGCGTGGTGGTATCAAGAAG	900
Db	993	GTTGAGTCAGGAGTCAGACTCTCCGATGCACGGTTCATGCGTGGAGTCCGCATCAAGAAG	1052
QY	901	CATGCTTCATCTCAACACAGATATTCGCGTGGCGACTCAACTGTTGGTCAATGGGCACGG	960
Db	1053	CATCGTGTATCTCGAGCAGTATCATCGGTTGGCACTCAACGGTTGGTCAATGGGCCAGG	1112
QY	961	ATAGAGATATGACTATCTCGGGGAGGATGTTCTATGTTGTGATGAGGTGTACAGCAAT	1020
Db	1113	ATCAGAAATCATGACATCTCTCGGTGAGGATGTTCTATGTCGAGCATGAGATCTATAGCAAT	1172
QY	1021	GCGGTGTTGTTCTCCACATAAAGAGATCAAGTCAAGCATCTCTCAAGCCTGAGATCGTC	1080
Db	1173	GGAGGATGTTTTCACACACAGGAGATCAATCAACATCTTCAAGCCAGATAGTG	1232
QY	1081	ATGTGA	1086
Db	1233	ATGTGA	1238
RESULT 8			
LOCUS	AP004127	142885 bp	DNA linear PLN 21-MAR-2002
DEFINITION	Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1,		
ACCESSION	AP004127		
VERSION	AP004127.1	GI:15408872	
KEYWORDS	Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,		
SOURCE	clone:P0005H10.		
ORGANISM	Oryza sativa (japonica cultivar-group)		
REFERENCE	Sasaki,T., Matsumoto,T. and Yamamoto,K.		
AUTHORS	Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC		
TITLE	clone:P0005H10		
JOURNAL	Published Only in Database (2001)		
REFERENCE	2 (bases 1 to 142885)		
AUTHORS	Sasaki,T., Matsumoto,T. and Yamamoto,K.		
TITLE	Direct Submission		
JOURNAL	Submitted (31-AUG-2001) Takuji Sasaki, National Institute of		
COMMENT	Agrobiological Sciences, Rice Genome Research Program; Kannondai		
	2-1-2, Tsukuba, Ibaraki 305-8602, Japan		
	(E-mail:tsasakienias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,		
	Tel:81-298-38-7441, Fax:81-298-38-7468)		
	Genes were predicted from the integrated results of the following:		
	GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as SplicePredictor		
	(October 1998 version). The genomic sequence was searched against		
	NCBI Nonredundant Protein database, nr		
	(ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at		
	RGF. Protein homologies of the coding regions were searched against		
	NCBI Nonredundant Protein databases using BLASTN 2.0 with the		
	identified cDNA sequences using BLASTN 2.0 with the		
	corresponding DDBJ accession no. and RGF clone ID.		
	A gene with identity or significant homology to a protein is		
	classified based on the protein name to indicate the homology level		
	such as same name, 'putative-' and '-like protein'. A gene without		
	significant homology to any protein but with EST homology (covering		
	almost the entire length of partial sequence) is classified as an		
	'unknown' protein. A gene predicted with a gene prediction program		
	is classified as a 'hypothetical' protein.		
	The orientation of the sequence is from SP6 to T7 of the PAC clone.		

Detailed information on overlap and assembly quality together with annotation of this entry is available at http://rgp.dna.affrc.go.jp/GenomeSeq.html.	
FEATURES	Location/Qualifiers
source	1. .142885
	/organism="Oryza sativa (japonica cultivar-group)"
	/cultivar="Nipponbare"
	/db_xref="taxon:39947"
	/chromosome="1"
	/clone="P0005H10"
gene	complement(join(3920. .3951,4840. .4981))
	/gene="P0005H10.1"
CDS	complement(join(3920. .3951,4840. .4981))
	/gene="P0005H10.1"
	/note="hypothetical protein"
	/codon_start=1
	/protein_id="BAB64264.1"
	/db_xref="GI:15408873"
	/translation="MRGGAAAGYAEAGSAGVVEEPRRVRVRKRREEVEAACTMA
	IGESWRTLIQYIF"
gene	join(6179. .6332,7620. .7781,8118. .8296,8558. .8683,
	8722. .8853,9433. .9540,9645. .9863,9939. .10119,10671. .10792)
	/gene="P0005H10.2"
CDS	join(6179. .6332,7620. .7781,8118. .8296,8558. .8683,
	8722. .8853,9433. .9540,9645. .9863,9939. .10119,10671. .10792)
	/gene="P0005H10.2"
	/note="contains EST C26260(C11984)
	similar to Arabidopsis thaliana chromosome 2, T517.4
	unknown protein"
	/codon_start=1
	/protein_id="BAB64265.1"
	/db_xref="GI:15408874"
	/translation="MSINFPVNVLEKCTEDILSLIKPVEGRNKRKIYAIQELADTI
	XSAGALRGASVRFPGSFVSQLYAKSGDLDSVELFNALNLPISKRKKQTLREVRRAL
	QKRJARHMEFIPNARVPLQVSNQYISDCISINSGRISKIFKYWINTLDDRFQ
	DMVLITCEPAIPLKEIYEGNIMEDISDKISLWEYGLPMDSDYLFGRAYNEKHLDN
	EVCSTNIERFRQNMQRNQSLSHLLASFFHKFRIDALSDKIVSTYGRLEIRIODN
	PRWMDKSLFVEDPEKPDNAARAVGSFODIVNAFNASNKVFSDAHLTDNRGL
	LSLLCTPVGSKLGRASRYNTLPAERGSGRSGSPHGYTTRQATVAYONNHQ
	AYNTQRQTVVHQNNQNVYAGRGQTGDNKLNQAGTSTMRSTLANLTITEHRNM
	HVTRRRNAL"
gene	complement(join(13608. .13812,14087. .14164,14524. .14646,
	15045. .15302,15753. .15912,16541. .16651,16998. .17268,
	17353. .17502,18222. .18929))
	/gene="P0005H10.3"
CDS	complement(join(13608. .13812,14087. .14164,14524. .14646,
	15045. .15302,15753. .15912,16541. .16651,16998. .17268,
	17353. .17502,18222. .18929))
	/gene="P0005H10.3"
	/codon_start=1
	/product="ankyrin-like protein"
	/protein_id="BAB64266.1"
	/db_xref="GI:15408875"
	/translation="MKPLCAHGADLLCGPRVLRALAAAFACAAFLVGVHWSDSPRL
	LFPSSTSSSSPLSLTSGSPSVAVSPNANLSDPSLIPTPAASTPASPATANASPPS
	LPFPPLPLRPPLPPARLGIVGEGAMRDFDVVGANDTDLAANTDEALPQETDAGP
	AVGSRVGRFPVPCESMREYIPCLDNEEIRRLPSTERGEFRCPCPKADKGLSLCV
	PAPKYKAPIPWPRSRDEVFNPHPTLVDKGGQNWISKADKFRPFGGQTQFTHG
	ANQYLDQISQWPDIAFGSHTRVALDVGGVAFGAYLSKRDVLTLSIAPKDVHENQI
	QFALSRGVPAAMAAFAFRLLYPSQAFDLIHCSRIRINWTHDDGILLVLRNMLRAGG
	YFAAQAQPVYHEEAQAQAKEMEDFARLCVELVKKEYIAMWRKPLANSYCMNRDP
	GVKPALCPDDNDPVDVWYNLKCISRLPENGSDGLTPFPWPAPRLMEPPKRLGVENDA
	VNSKKELEKAEKTFWDDIVGIVRFVKRKRKLNRVLDMRAGFGFPAALINKKLDGW
	VNNVVPVSEPNTPVYIDRGLGVADMCPEPTPTDYDLLHAFSLFKEQKRCNIS
	STILLEMDRLRPGRAYIRDLKQVVQDVKEITTAGMGRSIRMDTAEQVYASRVLKMD
	KPMV"
gene	Join(22002. .22041,22106. .22284)
	/gene="P0005H10.4"
CDS	Join(22002. .22041,22106. .22284)
	/gene="P0005H10.4"
	/note="hypothetical protein"
	/codon_start=1
	/protein_id="BAB64267.1"

/db_xref="GI:15408876"
/translation="MASNRIGWLAARAAHEVHRHSVOGRIINERAPKPVHGCTLVTRSH
REPCLPDRVMCCAKWFLVEYFRD"
Join(24898..25177,26029..26102,26187..26271,26369..26494,
26596..26832,26926..27162,27776..27819,28301..28483,
29186..29276,29798..29841,29924..30046)
/gene="P0005H10.5"
CDS
Join(24898..25177,26029..26102,26187..26271,26369..26494,
26596..26832,26926..27162,27776..27819,28301..28483,
29186..29276,29798..29841,29924..30046)
/gene="P0005H10.5"
/note="contains ESTs"
D48362(G14517).AU063373(C61107).AU092957(C61107)
similar to Arabidopsis thaliana chromosome 3, T2Z16.240
unknown protein"
/codon_start=1
/protein_id="BAB64268.1"
/db_xref="GI:15408877"
/translation="MNGRGCGCGGVVVGAGGEQMEDEAGLRLARYSAGRKERY
VLMGQFVGPQRSPRLGVPVRLPAAAGDGDWRDVGCGGCFAMAISESKLITWG
SADMGQSVTAGKHEETPEAPLPDSVAIVRADAGWAHVAITDEGDVYTWKKECV
PTGRVISDSQSGTLEKDEROSAIADQVSRSOVSPTSSGAASGPSRGTDSTKR
RRLSTKHAPRESSSDENLSAPPCVTFNTGVKIMAAAGRHHTLALSDLGQVNGW
YGGQGLSGSRITYSPHPICIESALYSKDRPAAKNGKSAEQVLSKVMGNCVKA
IACGRHSAVDSGALLTFWGLYGCQCGNTDDVLPTCVSSILGVMQDIDAGLW
HTVCTSDVDYFSGNQFGQGTGSDQAEVPLKVDASLENKRNARAYSCGARHSAI
ITDEGEYFCWGNRYGOLGDSMDRNPVCSVPVDAYHPLNVCWWHTLVLAESPT"
complement(Join(32368..32451,32633..32707,32945..33160,
33225..33448,33872..34085,34872..35405))
/gene="P0005H10.6"
CDS
complement(Join(32368..32451,32633..32707,32945..33160,
33225..33448,33872..34085,34872..35405))
/gene="P0005H10.6"
/note="hypothetical protein"
similar to Arabidopsis thaliana chromosome 1, F28G4.8"
/codon_start=1
/protein_id="BAB64269.1"
/db_xref="GI:15408878"
/translation="MADPPSAATASPODOLAAAVSTPONPNPILLSPOIPSPST
VSDLSAISPOLDPSAGCGAMDVPDRDQADPCQAAAGACGPGGTHRGGSRL
ANVGQLOPYAAARVHSQVNFSGCGGQVQGGQQLLAARAAHLSQQLHQLQGQ
GNAARHAGQQAQVAFKQNGVYGGVGFNTANAAALQGMVGMVGMGNGSTIP
YNQRFHAROLRPOOTSQGTLSPOVDFGCGSLDKNSIDCISTALQOQHLVQ
ADGNPSEFPAAGSVKNTNATITAGTCFSTISAKTDAINTCISTALQOQHLVQ
LQQQSPRIASGQSOKSMNLTSOPGTPLSGGTMVGGASOGAEVTLNQLLKKRQID
LYSOYDPLGKVDPEVDLLIADDFIDSVTAFACITLAKHRKSSVLEAKDVLHLG"
35814..36578
/gene="P0005H10.7"
CDS
35814..36578
/gene="P0005H10.7"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB64270.1"
/db_xref="GI:15408879"
/translation="MLSGRLEKTLTRDGNHSLPHVADRCORAKHCTLAWPCLSLPP
QDROSRLRFTDPRDPGSESTRSAQALAFALCRGSDARSPLSLATLSAAG
GNAHARVAYGDENTARLUDGTVSLARVLAFAVARLSYSGAEFRVDVA
ASLPARTFPFSLPLAKLVITYANRHPYSAYELPARGISLAFRCQFLLLYPII
TLSSRAHDFQFSSGRMAPRADAEKRRLPC"
Join(37250..37447,37511..37610,38449..38484,38714..38851,
39033..39261,39371..39432,39628..39707,40027..40268)
/gene="P0005H10.8"
CDS
Join(37250..37447,37511..37610,38449..38484,38714..38851,
39033..39261,39371..39432,39628..39707,40027..40268)
/gene="P0005H10.8"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB64271.1"
/db_xref="GI:15408880"
/translation="MBGLIPIFRVAVQYKGGHVSLSMDISDRPSRYSVLLPGDSV
GRHADDTOPILRQASGSEEDAVNCAAGASGSDNLRVPLICISADSEFVY
MGSVHNDLSPLSLHGRWHRGCKEGEGDMGRGGRFARQTAAGKREGGELI
YAGSHSSPMPCNRTAFELPPLSLSLSTVTAASILGSHPPPTALESPPVPHLGDPSD

SRSPTAIGVKLPRKPHVNLTWDELVKBATPSSCHREAIHHVTKNDPRDQPKSL
AHLGPVTRNKRISKSVQISVPESRLTERGAKIRSAGFFDASGLAHGISYMRD
YNGQRNASPWLACIYSLDR"
Join(41018..41128,41616..41678,41769..42008,42084..42755)
/gene="P0005H10.9"
CDS
Join(41018..41128,41616..41678,41769..42008,42084..42755)
/gene="P0005H10.9"
/note="contains ESTs"
/product="putative GDP-mannose pyrophosphorylase"
/codon_start=1
/protein_id="BAB64272.1"
/db_xref="GI:15408881"
Query Match 60.1%; Score 653; DB 8; Length 142885;
Best Local Similarity 81.3%; Pred. No. 5,4e-192;
Matches 804; Conservative 0; Mismatches 110; Indels 75; Gaps 1;
QY 173 AGCTAATGATTAATCTTGAAGACATTTGAGGATAGACTTGGCATCACAAATTCATGCT 232
DB 41767 AGCTAATGCTCAATTTCTGAAGACATTTGAGGATAGACTTGGCATCACAAATTCATGCT 41826
QY 233 CCCAAGAGAGCTGAGCCCTTAGGAACGCTGGCCCTCTTGTCTTAGCAAGGACAAAGCTTG 292
DB 41827 CCCAAGAGAGCTGAGCCCTTAGGAACGCTGGCCCTCTTGTCTTAGCAAGGACAAAGCTTG 41886
QY 293 CGGATGGATCTGCCAGCCATTTCTTGTCTCAACAGCTGATGTCATTAAGCCAATPACCCAT 352
DB 41887 TGGATGGATCTGGTGAGCCATTTCTTGTCTCAACAGCTGATGTCATTAAGCCAATPACCCAT 41946
QY 353 TTGCTCAACTCAATCAAAATTTCAAGGTCATGCTGTGTGAGGCAACAAATATATGTCACCTA 412
DB 41947 TTGCTGAGCTCATAAATTTCAAGGACCATGCTGTGTGAGGCAACAAATATATGTCACCTA 42006
QY 413 ----- 412
DB 42007 AGCTAATGCTTCTTTGACACCTCTCTGAAATTTGATAAATGCTGTGTATGCTGATACTGA 42066
QY 413 -----AGGTGGATGAACCATCAAAATACGGTGTGTGTGTGTATGCGAGGAGG 457
DB 42067 CTATTATTTCAATATAGGTGGAGCAACCATCAAAATATGCTGTGTGTGTATGCGAGGAGG 42126
QY 458 CAACCTGCGAGGCTGCAAGCTTTGAGAGAAAGGCAAAATATTTGCTGGCTGGCTAAAGATCA 517
DB 42127 TCATCGAATGTTGAGAAAATTTGTTGAGAAAACCAAAATATTTGTAGGCAACAAAGATCA 42186
QY 518 ATGCTGGGATTTACTTACTGAAACCATCTCTCTGTGACCCGATTTGACTGAGGCCACAT 577
DB 42187 ATGCGGCAATTTACTTTGCAATTCATCTCTGCGACCGCATGAGCTGAAAGCCAACTT 42246
QY 578 CAATTGAGAAAGAGGTCTTCCCTCAAAATTCGAGCTGATCAACAGCTCTTATGCAATGCTCC 637
DB 42247 CAATTGAGAAAGAGGTCTTCCCTCGAATTCGATCTGATGGAAGAGCTTTGCTCTGCTGCTC 42306
QY 638 TTCCAGGTTTTTGGATGGATTTGGCCAAACCAAGGATTTACATTTACAGGCTTGGCGCTTT 697
DB 42307 TTCCAGGTTTTTGGATGGATTTGGCCAAACCAAGGATTTACATTTACAGGCTTGGCGCTTT 42366
QY 698 ATCTAGACGATTTAGAGAAATACAGCTGCCAGCTAGCTACTGAGGACACATGTTGTTG 757
DB 42367 ATCTAGACGATTTAGAGAAATACAGCTGCCAGCTAGCTACTGAGGACACATGTTGTTG 42426
QY 758 GCAATGCTGTTGATGAGAGCGCAAGATTTGGAAGAGTTGTTGCTGATTGGTCTGCTGATG 817
DB 42427 GCAATGCTGTTGATGAGAGCGCGCAAGATTTGGAAGAGTTGTTGCTGATTGGTCTGCTGATG 42486
QY 818 TCGCCATTCGACCTGGCTGTGTTGTCGAGACGCGCTGAGCGCTTCCGCTGCACTGTCA 877
DB 42487 TTGCTATCGGCTCGTGTGATGGTGTGGAGGATGTTGTGGAGGCTTCCCGTTGACAGGTGA 42546
QY 878 TCGCGCGCTGCTGATGCAAGACATGCTTTCGATCTCAAAACAGACATTCGGCTGGCTGCA 937
DB 42547 TCGCGCGCTGCAATTAAGAGCATGCTTGCATATCAAAACAGACATTTATTTGCTGCTGCACT 42606

DEFINITION Oryza sativa (japonica cultivar-group) chromosome 8 clone P0682A06,
*** SEQUENCING IN PROGRESS ***, in ordered pieces.

ACCESSION AP004705

VERSION AP004705.1 GI:18447964

KEYWORDS HTG; HTGS_PHASE2.

SOURCE Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,
clone:P0682A06

ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.

REFERENCE 1

Sasaki, T., Matsumoto, T. and Yamamoto, K.

Oryza sativa nipponbare(GA3) genomic DNA, chromosome 8, PAC

clone:P0682A06

Published Only in Database (2002)

2 (bases 1 to 142024)

Sasaki, T., Matsumoto, T. and Yamamoto, K.

Direct Submission

Submitted (30-JAN-2002) Takuji Sasaki, National Institute of

Agronomical Sciences, Rice Genome Research Program; Kannondai

2-1-2, Tsukuba, Ibaraki 305-8602, Japan

(E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,

Tel:81-298-38-7441, Fax:81-298-38-7468)

NOTE: It currently consists of 1 contigs. gaps between the contigs
are represented as runs of N. The order of the pieces is believed
to be correct as given, however the sizes of the gaps between them
are based on estimates that have provided by the submitter. This
sequence will be replaced by the finished sequence as soon as it is
available and the accession number will be preserved.

* NOTE: this is a 'working draft' sequence.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

FEATURES

source

1..142024

/organism="Oryza sativa (japonica cultivar-group)"

/cultivar="Nipponbare"

/db_xref="taxon:39947"

/chromosome="8"

/clone="P0682A06"

BASE COUNT 39960 a 30290 c 31009 g 40765 t

ORIGIN

Query Match

Best Local Similarity 60.0%; Score 651.4; DB 2; Length 142024;

Matches 803; Conservative 0; Mismatches 111; Indels 75; Gaps 1;

Qy 173 AGGTAATGATTAATTTCTTGAAGGACTTTGAGGATAAGCTTGGCATCACAATCACGCTGT 232

Db 64862 AGGTAATGCTCAATTTCTCTGAAGGACTTTGAGGATAAGCTTGGCATCACAATCACGCTGT 64921

Qy 233 CCCAAGACACTGAGCCCTTAGGAACCGCTGGCCCTTGTCTAGCAAGGGACAGCTTG 292

Db 64922 CCCAAGACACTGAGCCCTTAGGAACCGCTGGCCCTTGTCTAGCAAGGGACAGCTTG 64981

Qy 293 CGGATGGATCTGGCCAGCCATTTCTTGTCTCAACAGTGTCTCATTAACGGAATACCCAT 352

Db 64982 TGGATGGATCTGGTGAGCCATCTTTGTCTCAACAGTGTCTCATTAACGGAATACCCAT 65041

Qy 353 TTGCTGAACATCATCAAAATTTTCAAAAGTGTGAGGAGGAGCAACAATATGGGTCACTA 412

Db 65042 TTGCTGAGCTCATAAATTTTCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 65101

Qy 413 ----- 412

Db 65102 AGGTAGATCCTTTCTTGACCACCTTTCTTGGATTATAAATCTGTTGTTGTATACTCA 65161

Qy 413 -----AGGTGATGATCAACCATCAAAATACGGTTGTGGTTATGGAGGAGG 457

Db 65162 CTATATTTTCAATGTAGTGGACGACCAACCATCAAAATATGGTTGTGGTTATGGAGGAGG 65221

Qy 458 CAACTGGCAGGGTGGAAAGGTTTGTGAGAACCCCAAAATATTTTGTGGGTAAACAAGATCA 517

Db 65222 CCACCTGGAATGGTGGAAAAATTTGTTGAGAAACCAAAAAATATTTGTAGGCACCAAGATCA 65281

Qy 518 ATGCTGGGATTTACTTACTGAACCCATCTGTCTTGACCGCATTTGAGCTGAGGCGCAACAT 577

Db 65282 ATCGAGGAATTTACTTGTGAACCCGCTGTCTCTGGACCGCATCGAGCTAAGCCACTT 65341

Qy 578 CAATGAGAAAGAGGCTTCCCTCAAAATGTCAGCTGATCAACAGCTCTATATGCAATGGTCC 637

Db 65342 CAATGAGAAAGAGGCTTCCCTCGAATCTCAGCTGATGCAAAAGCTTTTGTCTATGGTCC 65401

Qy 638 TTCCAGGTTTTTGATGGATGTTGGTCAGCCTAGGGACTACATTTACTTGGCTTGGCTCTTT 697

Db 65402 TTCCAGGTTTTTGATGGATGTTGGCCAACCAAGGATTACATTTACAGGCTTGGCCCTTT 65461

Qy 698 ATCTAGACTCGATTAGGAAGAAATACGTCGCAAGCTAGCTTACTTGGAGCACATTTGTTG 757

Db 65462 ATCTGATTCACATTAGGAAGATCAGCCAAACAGGTTAGCCACTGGAGCACACATTTGTTG 65521

Qy 758 GCAATGTGCTGGTCATGAGAGCGCCCAAGATTGGAGAAGGTTGTCTGATTTGGTCTCTGATG 817

Db 65522 GGAATGTGCTAGTTTACAGAGATGCGAAGATTGGCGAAGGCTGTCTGATTTGGTCTCTGATG 65581

Qy 818 TCGCCATTGGACCTGGGTGTTGTGAGGACGCGTGAGGCTTCCCGCTGCACTGTCA 877

Db 65582 TTGCTATTGGTCTCGATGCGTCTGAGGATGGTGTGAGGCTTCCCGTTTGCACGCTGA 65641

Qy 878 TGGCGGCGTGGCTATCAAGAAGCATCTTGTGCTCTCAACAGCATTTATCGCTGGCGGCT 937

Db 65642 TGGTGGCGTGGCTATTAAAGAGCATCTTGCATATCAACAGCATTTATTTGGATGGCACT 65701

Qy 938 CAACTGTGTGCTCAATGGGACGATAGAGAATATGACTATCTCTGGGGAGGATGTTCTATG 997

Db 65702 CAACTGTGGACAATGGGACGATAGAGAATATGACTATCTCTAGGAGGACGCTACATG 65761

Qy 998 TGTGTGATGAGGTGATCAGCAATGGCGGTGTTGTTCTCCCATATAAGAGATCAAGTCAA 1057

Db 65762 TAGGTGATGAGGTCTATACCAACGGCGGTGTTATTTCTCCGACAAAGAGATCAATCAA 65821

Qy 1058 GCATTTCTGAAGCCTGAGATGCTCATGTGA 1086

Db 65822 GCATCTGAAACCTGAGATTGTCTATGTGA 65850

RESULT 11

A97607

LOCUS A97607 Sequence 1 from Patent WO9915674.

ACCESSION A97607

VERSION A97607.1 GI:6780908

KEYWORDS

SOURCE potato.

ORGANISM Solanum tuberosum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Asteridae; euasterids I; Solanales; Solanaceae; Solanum.

REFERENCE 1 (bases 1 to 1478)

Keller, R. and Kossman, J.

AERIAL PARTS

TRANSGENIC PLANTS HAVING EARLY-RIPENING AND COMPLETELY WITHERING

Patent: WO 9915674-A 1 01-APR-1999;

MAX PLANCK GESELLSCHAFT (DE); KELLER RUTH (DE)

FEATURES

Location/Qualifiers

1..1478

/organism="Solanum tuberosum"

/db_xref="taxon:4113"

/tissue_type="BLATTGEWEBE"

/clone_lib="CDNA-BIBLIOTHEK AUS BLATTGEWEBE"

/clone_lib="(B) CLON(E): PGMFASE"

/notes="unnamed protein product"

/codon_start=1

/protein_id="CAB70521.1"

CDS

Db	256	CTCAAGGCTGTGGAGTAACCGAAGTTGTACTGGCTATTAATACTACCAACCTGAGGTGATG	315
QY	181	ATTAATTTCTTGAAGGACTTTGAGGATAAGCTTGATGCATCACAAATACATGCTCCCAAGAG	240
Db	316	CTGAACTTTCTTGAAGAATTTGAGGCAAGCCTTGGAATCAAGATCACCTGTTCTCAAGAA	375
QY	241	ACTGAGCCCTTAGGAACCGCTGGCCCTCTTGTCTCTAGCAAGGACAAGCTGCGGATGGA	300
Db	376	ACTGAACACACTGGCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	435
QY	301	TCTGCCAGCAGCTCTTCTGCTCAACAGTGTATGTCATAAGGGAATACCAATTTGCTGAA	360
Db	436	TCTGGTGAACCAATTTTCTTCAACAGTGTATGTCATAAGGGAATACCAATTTGCTGAA	495
QY	361	CTCATCAAAATTTCAACAGTGTATGCTGCTGAGGCAACAAATTTATGCTACTAAGTGGAT	420
Db	496	ATGATTTCAATTTCCACAAATCCCATGCTGCTGAGGCTTCTTGTGATGCTGACCAAGTGGAT	555
QY	421	GAACCATCAAAATACCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	480
Db	556	GAGCCTTCTTAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	615
QY	481	GTGAGAGCCAAATATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	540
Db	616	GTGAGAGCCAAATATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	675
QY	541	CCATCTGCTCTGAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	600
Db	676	CCTTCTGCTTACAGAGATTTCAATTTAGGCAACATCAATTTAGGCAAGGATTTTTCCTCA	735
QY	601	CAAAATGAGCTGATCAACAGCTCTATCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTG	660
Db	736	AAAATTTGAGCAGAGAGAAATGATCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	795
QY	661	GCTGAGCCTAGGACTACATTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	720
Db	796	GGCCAAACCAAGAGATTTACTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	855
QY	721	TCAGTGTGCAAGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	780
Db	856	TCTTCACTTAATTTGGCTTACAGGACCAACATTTGCGGAAATGTCATGCTGATGAATCT	915
QY	781	GCCAAAGATTGGAGAGGTTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	840
Db	916	GCCAAAGATTGGAGAGGTTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	975
QY	841	GTGAGGAGCGGTGAGGCTTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	900
Db	976	ATTGAGTCTGGAGTTAGACTCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1035
QY	901	CATGCTGCATCTCAACAGCATTTATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	960
Db	1036	CATGCTGCATCTCAACAGCATTTATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1095
QY	961	ATAGAGAAATGACTATCTCTGGGAGGATGTTCTATGCTGCTGCTGCTGCTGCTGCTGCTG	1020
Db	1096	GTTGAGACATGACCAATTTCTGGGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1155
QY	1021	GGCGGTGTTGTTCTCCCAATAAAGAGATCAAGTCAAGCAATTTCTGAAGCCTGAGATCGTC	1080
Db	1156	GGAGGTGATGTTTGGCCCAAGAGATCAAAATCCAGTATATTTGAAACCTGAAATCGTG	1215
QY	1081	ATGTGA 1086	
Db	1216	ATGTGA 1221	
RESULT 13			
AB066279			
LOCUS	1471 bp	mRNA	linear
DEFINITION	Nicotiana tabacum GMPase mRNA for GDP-D-mannose pyrophosphorylase, complete cds.		
ACCESSION	AB066279		

AB066279.1	GI:14971012	Nicotiana tabacum cell_line:BY-2 cDNA to mRNA.
KEYWORDS	Nicotiana tabacum	
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.	
REFERENCE	Tabata,K., Takaoka,T., Tateishi,Y. and Esaka,M.	
AUTHORS	Gene expression of ascorbic acid-biosynthetic and -oxidizing enzymes and ascorbic acid content in tobacco	
TITLE	Unpublished	
JOURNAL	2 (bases 1 to 1471)	
REFERENCE	Esaka,M., Takaoka,T., Tateishi,Y. and Tabata,K.	
AUTHORS	Direct Submission	
TITLE	Submitted (18-JUL-2001) Muneharu Esaka, Hiroshima University,	
JOURNAL	Faculty of Applied Biological Science, Kagamiyama, Higashi-Hiroshima, Hiroshima 739-8528, Japan	
FEATURES	(E-mail:mesaka@hiroshima-u.ac.jp, Tel:81-824-24-7927, Fax:81-824-24-7927)	
Source	Location/Qualifiers	
gene	1..1471	
CDS	/organism="Nicotiana tabacum"	
	/db_xref="taxon:4097"	
	/cell_line="BY-2"	
	/note="cultured"	
	1..1471	
	/gene="GMPase"	
	220..1305	
	/gene="GMPase"	
	/GC_number="2,7,7,22"	
	/function="ascorbic acid biosynthetic enzyme/carbohydrate biosynthetic enzyme"	
	/codon_start=1	
	/product="GDP-D-mannose pyrophosphorylase"	
	/protein_id="BAB62108.1"	
	/db_xref="GI:14971013"	
	/translation="MKALILVGGFGRPLRLTSLVPKPLVEFANKPMILHOIEALKAV	
	GTEVVLAINYOPEEMLNELKFEANLGIKTCSETEPLGTAGPLALARDKLIDDSG	
	EPPFVLNSDVISSEYPKEMIAFKSHGGASLMVTVDFPSKYGVVMEBSQTQVERF	
	VBKPLFVGNKINAGVILNPSVLDRIQRPTSEIEKVEFPKIAAEKRLYAMVLPFWM	
	DIGQPRDIYPLRLYLDLSKNSPKLASGSHVGNVIVDESAKIEGCLIGDPDAIG	
	SGCVIESGVRLSRCTVMGRVIRIKKHACISSIIGHSTVQMWARVENMTILGEDVHVC	
	DELYSNGGVLPKHEIKLSILKPEIVM"	
BASE COUNT	407 a	268 c 350 g 446 t
ORIGIN		
Query Match	58.7%	Score 638; DB 8; Length 1471;
Best Local Similarity	74.2%	Pred. No. 1.3e-187;
Matches	806; Conservative	0; Mismatches 280; Indels 0; Gaps 0;
QY	1	ATGAAGCCCTCATTTCTTTCGGGGTTTCGGAACCGCCTTCGGCCTTTGACTCTGAGC 60
Db	220	ATGAAGCCGCTTATTTCTTTCGGGGTTTCGGAACCGCCTTCGGCCTTTGACTCTGAGC 279
QY	61	TTCCCGAAACCCCTCGTGGATTTTGCAACAAGCCCATGATCTGCACCATGGAAGCT 120
Db	280	GTCCCAAGCCATAGTTGAATTTGCCAACAACCCCATGATTTGCATCAGATTGAGGCT 339
QY	121	TTGAAAGAAGTTGGGTCACAGAGGTGTTTGGCTATCAACTATCCCGCAGAGGTAATG 180
Db	340	CTCAAGCAGTCGGAGTAAGTTGACTGGCTATTATTTACCAACCTGAGGAGATG 399
QY	181	ATTAATTTCTTGAAGGACTTTGAGGATAAGCTTGGCATCAAAATTTACATGCTCCCAAGAG 240
Db	400	CTGAACCTTCTTGAAGAATTTGAGGCAACCTTGGAAATCAAGATCACTGTCTTCAAGAA 459
QY	241	ACTGAGCCCTTAGAACCCTGGCCCTTCTGCTCTACAGGAGGACAGCTTGGCGATGGA 300
Db	460	ACTGAACCACTTGGCACTGCGAGGTCCCTTGTGCTTAGAGATAAGATGATGATGAC 519
QY	301	TCTGGCCAGCCATTTCTTGTGCTTCAACAGTGATGCTCATAGCGAATACCCATTTGCTGAA 360

DD	32691	TAGGTACACAGATCAAGCGCTGGGATTTATCTTCTGACACCATCTGTTCTTGATAAGATTG	32750
QY	563	AGCTAGGCGCAAGATCAATTTAGAAAGAGAGTCTTCCCTCAAAATTTGACGCTGATCAACAGC	622
DB	32751	ACGTAAACAGCGGACTTCAATCGGAAAGACAGACTTTCCTTAAGATTGCGAGCAGCGCAAGGCG	32810
QY	623	TCTATGCAATGGTCTTCCAGGTTTGGATGATGATGTTGGTCAGCCVAGGCGCATACATTAA	682
DB	32811	TCTATGCTATGGTGTCTACCAAGGTTTGGATGGACATTTGGCAACCCCGTGACTACATAAA	32870
QY	683	CTGCGCTTGGCTTCTATCTAGACTCCATTAGGAGAAATACAGCTGCCAAGCTAGCTACTG	742
DB	32871	CGGGTTTGAGACTCTACTTAGACTCCCTTAGGAGAAATCTCTCGCAATTAACAGTG	32930
QY	743	GAGCAGATGTTGTGCAATGTGCTGGTGCATGAGAGCGCAAGATTGGAGAGGTTGTC	802
DB	32931	GCCACACATAGTTGGAAATGTTCTGTTTCACGAACCCGCTACAATTTGGGAGGATGTT	32990
QY	803	TGATTGGTCTGATGTCGCATTTGACCTGGGTGTTGTTGGAGGCGGCTGAGGCTTTT	862
DB	32991	TGATTGGACAGACGCTGTCATTTGTCAGGCTGTCATTGTTGAGTCAGGAGTCAGACTCT	33050
QY	863	CCGCTGCACTGTCATCGCGCGCTGCTATCAAGAGCATGCTTCATCTCAACAGCA	922
DB	33051	CCGATGACAGGTCATCGCTGGAGTCGCGATCAAGAGCATGCGTGTATCTCAGCAGTA	33110
QY	923	TTATCGCTGGCACTCAACTGTTGGTCAATGGCACGGATAGAGATATGACTATCTCTG	982
DB	33111	TCATCGGTGGCACTCAACGGTGGTCAATGGCCAGGATCGACACATGACCATCTCTCG	33170
QY	983	GGGAGATGTTCATGTGTGATGAGGTGTACAGCAATGGCGGTGTTGTTCTCCCATATA	1042
DB	33171	GTGAGGATGTTGATGAGGATGAGATCATAGCAATGGAGGAGTGTGTTTCCACACA	33230
QY	1043	ACAGATCAACTCAACCATTCGAACCCTCAGATCGTCATGCA	1086
DB	33231	AGGAGATCAATCAACATCTTTGAAGCCAGAGATAGTGATGA	33274
RESULT	15		
ATT22E16			
LOCUS	ATT22E16	103240 bp	DNA linear
DEFINITION	Arabidopsis thaliana DNA chromosome 3, BAC clone T22E16.	PLN 23-FEB-2000	
ACCESSION	AL132975		
VERSION	AL132975.1	GI:6434228	
KEYWORDS	Arabidopsis thaliana.		
SOURCE	Arabidopsis thaliana		
ORGANISM	Arabidopsis thaliana		
REFERENCE			
AUTHORS	Benes,V., Wurmbach,E., Drzonek,H., Ansorge,W., Mewes,H.W., Lemcke,K., Mayer,K.F.X., Quetier,F. and Salanoubat,M.		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 103240)		
AUTHORS	EU Arabidopsis sequencing project.		
JOURNAL	Submitted (23-FEB-2000) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail: lemcke@mips.biochem.mpg.de, mayermips.biochem.mpg.de Project Coordinator: Marcel Salanoubat and Francis Quetier, Groupement d'Interet Public, Centre National de Sequencage - GENOSCOPE, 2 rue Gaston Cromieux, BP191, 91006 Evry Cedex, France; http://www.genoscope.cns.fr		
COMMENT	Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/tha/.		
FEATURES			
source	1..103240		
	/organism="Arabidopsis thaliana"		
	/variety="Columbia"		
	/db_xref="taxon:3702"		

misc_feature	/chromosome="3"	
gene	1..8852	
	/note="Overlap to BAC T26112 (EMBL:AL132954)"	
	992..3164	
	/gene="T22E16.10"	
CDS	join(992..1181,2134..3164)	
	/gene="T22E16.10"	
	/note="similarity to predicted proteins, Arabidopsis thaliana"	
	/codon_start=1	
	/product="putative protein"	
	/protein_id="CAB75893.1"	
	/db_xref="GI:7076778"	
	/translation="MGPIKTIKKKRAEKVDRNVLLAATAATAASASAAAALNNDDDDSSSSOSLMDWDFSRRIYGGSTDPKTFESVFKISRTFDYICSLVADFTAKPANESDSNGPLSLNDRVAVLRIGSGESLIVIGETFCMNOSTVSOITWRVSEMERAIHLSNPKLDEIKSFEKISGLPNCGAIDITHIYVNLPAVSPKRWLDGKFNSTLQAVYDPMRFLDYJAGWPGSUNDVYLNKSGFYKLVKRGKNGKLPSESTREYIIVGDSGFLLPWLTPYCGKPTSLPOTEFNKRHSATKAQMAISKLKDRWRIINGVMMPDRNLRPRIIFYVCLLHNTIIDMEDQTLDDQLSQQHDHMYRQSRKLADEASSVLRDELSDLQCGKNSSA"	
exon	992..1181	
	/gene="T22E16.10"	
	/number=1	
intron	1182..2133	
	/gene="T22E16.10"	
	/number=1	
exon	2134..3164	
	/gene="T22E16.10"	
	/number=2	
gene	3660..5330	
	/gene="T22E16.20"	
CDS	complement(join(3660..3728,3821..4575,5038..5077,5262..5330))	
	/note="strong similarity to synaptic glycoprotein SC2 spliced variant, Homo sapiens, EMBL:AF038998"	
	/codon_start=1	
	/product="glycoprotein-like"	
	/protein_id="CAB75894.1"	
	/db_xref="GI:7076779"	
	/translation="MKVTVVSRSGRVILKAPLDLPDSATVADLQEAHRAKRPYPSRQRUTLPVTGSKDPVYLNKSKSCYCGNNNSLTVYFKDLGAQYSYRTLFFEEYLGPLLIVPVYFFVYKFGXGDCVLIHPQVTAMY YMCFFHYFKRIETFFVHRFSHATSPIGNVPRNCAVYVSGAYTAYVNHPLYTPVSDLMQKRGFGGLGVCOVAFYCHILK NLRDPSAGGYQIPRGLENIVTCANTTEIYQWLGNIATQITAGYVFLAVALIINTNWALGKHSRLKIFDKGDKPKYPRKRVILPPEL"	
exon	complement(3660..3728)	
	/number=1	
intron	complement(3729..3820)	
	/number=1	
exon	complement(3821..4575)	
	/number=2	
intron	complement(4576..5037)	
	/number=2	
exon	complement(5038..5077)	
	/number=3	
intron	complement(5078..5261)	
	/number=3	
exon	complement(5262..5330)	
	/number=4	
gene	10334..11350	
	/gene="T22E16.30"	
CDS	join(10334..11042,11136..11350)	
	/gene="T22E16.30"	
	/note="EMBL:AF155818 most likely is not full length, therefore N-terminal stretch"	
	/codon_start=1	
	/product="zinc finger protein OBP3"	
	/protein_id="CAB75895.1"	
	/db_xref="GI:7076780"	
	/translation="MOOGNHOLECVTTONPNNYIROLSSPPTSOVAGSSOARVNMVFRATARKVLPALPAALNCPDSTNFKCFVFNYSLTOPHFCKTCRRVTRGGLRN VFGVGGFRNRKRSKSKSTVYVYSTDNTTSTSSLSIRFSYSNFSKFSYIQIPEFNSN	

LPILPLOSQDYNSSNTGLDFGQTOISNMISGSSSGGILDWIRIPPSQQAQPPFL
 INTGLVQSSNALPPLLEGGVSATQTRNVAEENDQDRGDGVNLLSRNPLGNINI
 NSGRNEYTSWGNSSWTGFTSNNSTGLHSF"
 10334..11042
 /gene="T22E16.30"
 /number=1
 11043..11135
 /gene="T22E16.30"
 /number=1
 11136..11350
 /gene="T22E16.30"
 /number=2
 13994..15369
 /gene="T22E16.40"
 join(13994..14039,14159..14261,14513..14610,14695..14795,
 15214..15369)
 /gene="T22E16.40"
 /notes="Contains Ubiquitin-conjugating enzymes active site
 AA79-93"
 /codon_start=1
 /product="ubiquitin-conjugating enzyme UBC3"
 /protein_id="CAB75896.1"
 /db_xref="GI:7076781"
 /translation="MANQASLLLOKQKLDLKKPVDGFSAGLVDEKNVFWMSVIMG
 PPDTLYEGGFNAIMSPENYPVSPPTVTFSEMHPNVYSDGKVCISILPPGDDPH
 GYELASERWTPVHTVESIVLSIISMLSGPNDESPANVEAAKEWRDRAEFRKVSRCV
 RRSQEML"
 13994..14039
 /gene="T22E16.40"
 /number=1
 14040..14158
 /gene="T22E16.40"
 /number=1
 14159..14261
 /gene="T22E16.40"
 /number=2
 14262..14512
 /gene="T22E16.40"
 /number=2
 14513..14610
 /gene="T22E16.40"
 /number=3
 14611..14694
 /gene="T22E16.40"
 /number=3
 14695..14795
 /gene="T22E16.40"
 /number=4
 14796..15213
 /gene="T22E16.40"
 /number=4
 15214..15369
 /gene="T22E16.40"
 /number=5
 17259..18094
 /gene="T22E16.50"
 join(17259..17479,17564..17699,17867..18094)
 /gene="T22E16.50"
 /note="similarity to MtN24, Medicago truncatula,
 EMBL:WYI15290; Contains prokaryotic membrane lipoprotein
 lipid attachment site AA164-174"
 /codon_start=1
 /product="putative protein"
 /protein_id="CAB75897.1"
 /db_xref="GI:7076782"
 /translation="MRSPHAFRNGESPTRLDRHTFHSTVTAQKLRRRENSLILLRLAS
 FFSLSASAVFMLTNSRGSASPHWYDFDAFRFVANAIVALSVPFEMGTCVWESRET
 TLMPFAFWFDGHQVFSYLLSAGSAAALARTMRGGDTCTANKAFCLQSDVAIG
 LGAAFLFLAFSSCFGRVACFLITGSRFHLYS"
 17259..17479
 /gene="T22E16.50"
 /number=1
 17480..17563

/gene="T22E16.50"
 /number=1
 17564..17699
 /gene="T22E16.50"
 /number=2
 17700..17866
 /gene="T22E16.50"
 /number=2
 17867..18094
 /gene="T22E16.50"
 /number=3
 18283..21473
 /gene="T22E16.60"
 complement(join(18283..18465,18558..18686,18800..18955,
 19051..19166,19272..19452,19597..19695,19991..20084,
 20260..20384,20536..20703,20791..21084,21168..21473))
 /codon_start=1
 /product="methionyl-tRNA synthetase (AtcpMetRS)"
 /protein_id="CAB73898.1"
 /db_xref="GI:7076783"
 /translation="MAARINTSLHNALSLKPFNPPLNTKPFSTRNSRFRSKLPPY
 SOfSSGKRALYCTSSQESTVDEGETFVLTTPLYVYNAPPHMGSAVTTIAADSIARFQ
 Query Match 42.5%; Score 461.2; DB 8; Length 103240;
 Best Local Similarity 68.7%; Pred. No. 4.5e-132;
 Matches 692; Conservative 0; Mismatches 238; Indels 78; Gaps 1;
 QY 157 ATCAACTATCGCCAGAGGTAAATGATTCTTCTGAAGACATTTGAGGATAAGCTTGGC 216
 DB 100350 ATCAATTCAGCAACTGCTAGTAGTGTCCTCAATCTCAACAGCAGCTTGGGACCTCGG 100409
 QY 217 ATCAAAATACATGCTCCCAAGAGACTGAGCCCTTAGAACCCGCTGGCCCTCTTGCTCTA 276
 DB 100410 ATCAAGATCACCTGCTCAAGAGACTGAGCCTTAGGTACAGCTGCTCTTGGCTTTA 100469
 QY 277 GCAAGGGACAAGCTTCGGGATGATCGCCAGCCATCTTTTGTCTCAACAGTGTATGTC 336
 DB 100470 GCCAGGGACAACTGGTGTGCTCTGGCAACCATCTTGTCTCTAACAGCGATGTT 100529
 QY 337 ATAAGCGAATACCCATTTGCTGAACATCAAAATTTCAAAAGTGTCTGTTGGTGAGGCA 396
 DB 100530 ATTAGCGACTACCCGCTTGAAGAGATGATGCTTTTCAAAATGCTCTCAGGAGAGGCC 100589
 QY 397 ACAATATATGCTCAC-----
 DB 100590 TCTATAATGCTGACAAAGGTAAATATTTTGTATGAACCTTTAAGCATTTGTGTTGATG 100649
 QY 411 -----
 DB 100650 TTTTCTTACTCATTATATATAATTTGTTTAAAGGTAGATGAGCCTTCGAAATATGGA 100709
 QY 439 GTTGTGGTTATGAGGAGGCAACTGGCAGGTGGAAGGTTTGTGTAGAGACCCAAAATA 498
 DB 100710 GTTGTGTTATGGAAGAAGCCACAGGAGAGATTCGTCGAAAAACCCGAACTA 100769
 QY 499 TTTGTGGTAAACAAGATCAATGCTGGGATTTACTTACTGAACCATCTGCTCTGACCGC 558
 DB 100770 TTTGTGGTAAACAAGATCAATGCTGGGATTTATCTTCTAAATCCTTCAGTTCTTGATCGG 100829
 QY 559 ATTGAGCTGAGGCAACATCAATGAGAAAGAGTCTTCCCTCAAAATTTGCAAGCTGTATCAA 618
 DB 100830 ATTGAGCTAAGACCAACATCAATAGAGAAGAGATCTTCCCTCAATCGGAGAGCGGAA 100889
 QY 619 CAGCTCTATGCAATGGTCTCTCCAGGTTTTTGGATGATGTTTGGTCAAGCTAGGACTAC 678
 DB 100890 AAGCTCTATGCAATGCTCTCTCCGAGTTCTCGGATGACATTTGTCACACGAGACTAC 100949
 QY 679 ATTACTGGCTTGGCTCTTTATCTAGACTCGATTAGGAAGAAATCAGCTGCCAAGCTAGCT 738
 DB 100950 ATTACTGGTCTAAGACTTTTACTTTAGACTCTCTCTGAGGAAGAAATCTCTCTCAAAATGAGCC 101009
 QY 739 ACTGGAGCACATGTTCTTGTGCAATGCTGTCATGAGAGCCGAGATTTGGAGAGGT 798

Db	101010	ACTGGTCCACACATTTTGGGAAAGTCTTGTGGATGAAACCGCAGAGATGGTGAAGGA	101069
Qy	799	TGTCGTGATGGCTCCTCATCTTCGCCATTTGGACCTGGCTCTTCTCAGACAGCCGCTGAGG	858
Db	101070	TGTTTGATGGCCCTAACGTTGCTATAGGCCCTTGGCTGTGTTGTTGGTCCGGGGTTAGA	101129
Qy	859	CTTTCGCGCTGCACCTGCATGCGCGCGTGGGTATCAAGAAGCAGCTGTGCATCTCAAC	918
Db	101130	CTATCGCACCTGCACGGTAATGCGTGGGGTACACCTCAACAGATATGCGTCTATCPCGAGT	101189
Qy	919	AGCATATTCGGCTGGCAGCTCAACTGTTTGTCATATGGGCGAGGATAGAGAAATATGACTATC	978
Db	101190	AGTATCATTTGGTGGCACTCGACGGTCGGACAATGGGCTCGGAGAACATGTGCCATA	101249
Qy	979	CTGGGGGAGGATCTTCATCTGCTGTGATAGGCTCTACACGAATGGCGGTCTGTTCTCCCA	1038
Db	101250	CTTGGGAAAAATGTTACGTGTGCGGATGAGATATTTGTAATGGAGGGGTGTGTTGGAT	101309
Qy	1039	CATAAAGAGATCAAGTCAAGCAATCTGAAGCGCTGAGATCGTCAATGA	1086
Db	101310	AACAAGACATCAAAATCAGACATCTTTTGTAACCAACGATATAGTCAATGTGA	101357

Search completed: November 27, 2002, 03:38:11
Job time : 3091 secs

PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0132866.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139730.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140699.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142059.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144086.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 08-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157953.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158569.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.

PR	28-MAY-1999;	99US-0136782.	PR	04-AUG-1999;	99US-0147302.
PR	01-JUN-1999;	99US-0137222.	PR	05-AUG-1999;	99US-0147192.
PR	03-JUN-1999;	99US-0137528.	PR	05-AUG-1999;	99US-0147260.
PR	04-JUN-1999;	99US-0137502.	PR	06-AUG-1999;	99US-0147303.
PR	07-JUN-1999;	99US-0137724.	PR	06-AUG-1999;	99US-0147416.
PR	08-JUN-1999;	99US-0138094.	PR	09-AUG-1999;	99US-0147493.
PR	10-JUN-1999;	99US-0138540.	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	99US-0138847.	PR	10-AUG-1999;	99US-0148171.
PR	14-JUN-1999;	99US-0139119.	PR	11-AUG-1999;	99US-0148319.
PR	16-JUN-1999;	99US-0139452.	PR	12-AUG-1999;	99US-0148341.
PR	17-JUN-1999;	99US-0139453.	PR	13-AUG-1999;	99US-0148565.
PR	18-JUN-1999;	99US-0139454.	PR	13-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	99US-0139455.	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139456.	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139457.	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-0139458.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139459.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139455.	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	99US-0139460.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139461.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139462.	PR	25-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139463.	PR	26-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	99US-0139750.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139750.	PR	27-AUG-1999;	99US-0151065.
PR	21-JUN-1999;	99US-0139817.	PR	27-AUG-1999;	99US-0151066.
PR	22-JUN-1999;	99US-0139899.	PR	27-AUG-1999;	99US-0151080.
PR	23-JUN-1999;	99US-0140353.	PR	31-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	99US-0140354.	PR	31-AUG-1999;	99US-0151438.
PR	24-JUN-1999;	99US-0140354.	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	99US-0140698.	PR	07-SEP-1999;	99US-0152363.
PR	28-JUN-1999;	99US-0140823.	PR	10-SEP-1999;	99US-0153070.
PR	29-JUN-1999;	99US-0140991.	PR	13-SEP-1999;	99US-0153758.
PR	30-JUN-1999;	99US-0141287.	PR	13-SEP-1999;	99US-0153758.
PR	01-JUL-1999;	99US-0141842.	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0142154.	PR	16-SEP-1999;	99US-0154039.
PR	02-JUL-1999;	99US-0142055.	PR	20-SEP-1999;	99US-0154779.
PR	06-JUL-1999;	99US-0142390.	PR	22-SEP-1999;	99US-0155139.
PR	08-JUL-1999;	99US-0142803.	PR	23-SEP-1999;	99US-0155486.
PR	09-JUL-1999;	99US-0142920.	PR	24-SEP-1999;	99US-0155659.
PR	12-JUL-1999;	99US-0142977.	PR	28-SEP-1999;	99US-0156458.
PR	13-JUL-1999;	99US-0143542.	PR	29-SEP-1999;	99US-0156596.
PR	14-JUL-1999;	99US-0143624.	PR	04-OCT-1999;	99US-0157117.
PR	15-JUL-1999;	99US-0144005.	PR	05-OCT-1999;	99US-0157753.
PR	16-JUL-1999;	99US-0144085.	PR	06-OCT-1999;	99US-0157865.
PR	16-JUL-1999;	99US-0144086.	PR	07-OCT-1999;	99US-0158029.
PR	19-JUL-1999;	99US-0144325.	PR	08-OCT-1999;	99US-0158232.
PR	19-JUL-1999;	99US-0144325.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144331.	PR	13-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-0144333.	PR	13-OCT-1999;	99US-0159295.
PR	19-JUL-1999;	99US-0144334.	PR	14-OCT-1999;	99US-0159329.
PR	20-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159330.
PR	20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159331.
PR	20-JUL-1999;	99US-0144632.	PR	14-OCT-1999;	99US-0159637.
PR	20-JUL-1999;	99US-0144884.	PR	14-OCT-1999;	99US-0159638.
PR	21-JUL-1999;	99US-0144814.	PR	18-OCT-1999;	99US-0159584.
PR	21-JUL-1999;	99US-0145086.	PR	21-OCT-1999;	99US-0160741.
PR	21-JUL-1999;	99US-0145088.	PR	21-OCT-1999;	99US-0160767.
PR	22-JUL-1999;	99US-0145085.	PR	21-OCT-1999;	99US-0160768.
PR	22-JUL-1999;	99US-0145087.	PR	21-OCT-1999;	99US-0160770.
PR	22-JUL-1999;	99US-0145089.	PR	21-OCT-1999;	99US-0160814.
PR	22-JUL-1999;	99US-0145192.	PR	21-OCT-1999;	99US-0160815.
PR	23-JUL-1999;	99US-0145145.	PR	22-OCT-1999;	99US-0160980.
PR	23-JUL-1999;	99US-0145218.	PR	22-OCT-1999;	99US-0160981.
PR	23-JUL-1999;	99US-0145224.	PR	22-OCT-1999;	99US-0160989.
PR	26-JUL-1999;	99US-0145276.	PR	22-OCT-1999;	99US-0160989.
PR	27-JUL-1999;	99US-0145913.	PR	25-OCT-1999;	99US-0161404.
PR	27-JUL-1999;	99US-0145918.	PR	25-OCT-1999;	99US-0161405.
PR	27-JUL-1999;	99US-0145919.	PR	25-OCT-1999;	99US-0161406.
PR	28-JUL-1999;	99US-0145951.	PR	26-OCT-1999;	99US-0161359.
PR	02-AUG-1999;	99US-0146380.	PR	26-OCT-1999;	99US-0161360.
PR	02-AUG-1999;	99US-0146388.	PR	26-OCT-1999;	99US-0161361.
PR	02-AUG-1999;	99US-0146389.	PR	28-OCT-1999;	99US-0161920.
PR	03-AUG-1999;	99US-0147038.	PR	28-OCT-1999;	99US-0161992.
PR	04-AUG-1999;	99US-0147204.	PR	28-OCT-1999;	99US-0161993.
PR			PR	29-OCT-1999;	99US-0162142.

Query Match									
Best Local Similarity 63.6%; Score 690.8; DB 21; Length 1474;									
Matches 839; Conservative 0; Mismatches 247; Indels 0; Caps 0;									
QY	1	ATGAGGCGCTCATCTTGTGCGGGGTTTCGGACCGCCTTCGGCTTTGACTCTGAGC	60						
DB	155	ATGAGGCGCTCATCTTGTGCGGGGTTTCGGACCGCCTTCGGCTTTGACTCTGAGC	214						
QY	61	TTCCCAAGCCCTTGGATTTGCTATTAACCCATGATCTTATCAGATAGAGCT	274						
DB	215	TTCCCAAGCCCTTGGATTTGCTATTAACCCATGATCTTATCAGATAGAGCT	274						
QY	121	TTGAAAGAGTTGGGTGACAGAGGTGGTTTGGCTATCAACTATCCGCCAGAGTAAT	180						
DB	275	CTTAAGGCGCTTGCAGCTTGATGAAGTCTTGGCCATCAATTAACGAGAGGTGATG	334						
QY	181	ATTAATTTCTTGAGACATTTGAGGATAGCTTGGCATACCAATTACATGCTCCCAAG	240						
DB	335	CTGAAGCTTTGAGACATTTGAGACCAAGCTGGAATCAAAATCACTTCTCACAAG	394						
QY	241	ACGACGCGCTTAGCAACCGCTGCGCTTCGCTTACCAAGCAAGCTGCGCATGCA	300						
DB	395	ACCGAGCCACTAGTACCGCTGCTCTTGGCTTACGAGAGCAAAATTTGCTTGTGGA	454						
QY	301	TCGCGCAGCCATCTTTGCTCCTCAACAGTGTATCATTAAGCCGAATACCCATTTGCTGA	360						
DB	455	TCGCGCAGCCATCTTTGCTCCTCAACAGTGTATCATTAAGCCGAATACCCATTTGCTGA	514						
QY	361	CTCATCAAAATTTACAAAGTGTCAATGTTGGAGGCAACAATTAATGCTACAAAGTGGAT	420						
DB	515	ATGCTTCAGTTTACAAAATCTCAGCGTGGGGAAGCCCTCCATAATGTTAAACAAAGCTGAT	574						
QY	421	GAACCATCAAAATPACCGTCTGCTGCTTATCGAGGAGCGCACTGCGAGCTGGGAAGGTTT	480						
DB	575	GRACCTCGAAATATGGAGTGGTTGTTATGGAAGAAGCACTGGAGAGTGGGAAGTTT	634						
QY	481	GTTGAGAACCCAAAATATTTGGTGGTAAACAATCAATGCTGGGATTTACTTACTCAAC	540						
DB	635	GTTGAGAACCCAAAATATTTGGTGGTAAACAATCAATGCTGGGATTTACTTACTCAAC	634						
QY	541	CCATCTGCTGAGCGGATTTAGCTGAGGCGCAACATCAATTAAGAAAGAGTCTTCCCT	600						
DB	695	CCATCTGCTGAGCGGATTTAGCTGAGGCGCAACATCAATTAAGAAAGAGTCTTCCCT	754						
QY	601	CAAAATGACGCTGATCAACAGCTTATGCAATGGTCTCTCCAGGTTTGGATGGATGTT	660						
DB	755	AGATTCGACGAGCGCAAGGCTCTATGCTATGCTGCTACCGGTTTGGATGGACAT	814						
QY	661	GCTACGCTAGGACATACATTTACTGCTTCCGCTTCTTATCTAGACATCGATTAGCAAGAA	720						
DB	815	GGGCAACCCCGTACATACATACGCGTTTGGAGACTTACTTACTTCCCTTAGGAGAAA	874						
QY	721	TCAGCTGCCAAGCTAGCTACTGAGAGCAGATTTGTTGGCAATGCTGTGTCATGAGAGC	780						
DB	875	TCCTCTGCCAAAATTAACGAGTGGCCACACATAGTTGGCAATGCTGTGTCATGAGAGC	934						
QY	781	GCCAGATTTGGAGAGGTTTGTCTGATTTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCT	840						
DB	935	GCTACAATTTGGGAGAGGATTTGATTTGAGACAGAGCTTGGCAATGCTGTGTCATGAGAGC	994						
QY	841	GTCGAGGAGCGCTGAGGCTTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	900						
DB	995	GTTGAGTACAGAGTACAGCTCTCCGATGACAGGCTGATGCTGCTGCTGCTGCTGCTGCT	1054						
QY	901	CATGCTTCATCTCAACAGCAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	960						
DB	1055	CATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1114						
QY	961	ATAGAGCAATATGACTATCTGCTGGGAGGATGTTTCATGCTGTGATGAGGTGTACAGCAAT	1020						
DB	1115	ATCGAGCAATATGACTATCTGCTGGGAGGATGTTTCATGCTGTGATGAGGTGTACAGCAAT	1174						

QY	1021	GGCGGTGTTCTTCCACATATAGAGATCAAGTCAAGCATCTTCAAGCCTTGAGATCGTC	1080
DB	1175	GGAGAGTGTGTTTGGCACACACAGAGATCAATCAATCTTGAACCCAGAGATGTC	1234
QY	1081	ATGTCA 1086	
DB	1235	ATGTCA 1240	
RESULT 4			
AC38731			
ID	AAC38731	standard; DNA; 1480 BP.	
XX	AAC38731;		
XX	17-OCT-2000	(first entry)	
XX	Arabidopsis thaliana	DNA fragment SEQ ID NO: 22029.	
DE	Hybridisation assay; genetic mapping; gene expression control;		
KW	protein identification; signal transduction pathway;		
KW	metabolic pathway; promoter; termination sequence; ss.		
XX	Arabidopsis thaliana.		
OS	EP1033405-A2.		
PN	06-SEP-2000.		
PD	25-FEB-2000; 2000EP-0301439.		
XX	99US-0121825.		
XX	99US-0123180.		
PR	99US-0123548.		
PR	99US-0125788.		
DR	99US-0126364.		
PR	99US-0126785.		
PR	99US-0127462.		
PR	99US-0128234.		
PR	99US-0128714.		
PR	99US-0129845.		
PR	99US-0130077.		
PR	99US-0130449.		
PR	99US-0130510.		
PR	99US-0130891.		
PR	99US-0131449.		
PR	99US-0132048.		
PR	99US-0132407.		
PR	99US-0132484.		
PR	99US-0132485.		
PR	99US-0132486.		
PR	99US-0132487.		
PR	99US-0132863.		
PR	99US-0134256.		
PR	99US-0134218.		
PR	99US-0134219.		
PR	99US-0134221.		
PR	99US-0134370.		
PR	99US-0134768.		
PR	99US-0134941.		
PR	99US-0135124.		
PR	99US-0135353.		
PR	99US-0135629.		
PR	99US-0136021.		
PR	99US-0136392.		
PR	99US-0136782.		
PR	99US-0137222.		
PR	99US-0137528.		
PR	99US-0137502.		
PR	99US-0137724.		
PR	99US-0138094.		
PR	99US-0138540.		
PR	99US-0138847.		

PR 14-JUN-1999; 99US-01391119.
PR 16-JUN-1999; 99US-0139432.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139859.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.

PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0156559.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 63.6%; Score 690.8; DB 21; Length 1480;
Best Local Similarity 77.3%; Pred. No. 2.7e-210;
Matches 839; Conservative 0; Mismatches 247; Indels 0; Gaps 0;

QY 1 ATCAAGGCCCTCATTTCTTCGGGGGTTTCGGAACCCGCTTCGGCCTTTTGACTCTGAGC 60
DB 155 ATGAGGCGCATCTCTTGTGGAGGCTTCGGCACTCGCTTGAGACCATGACTCTCAGT 214

```
QY 61 TTCCCGAAGCCCTCTGCGATTTTGCAGAACCCCAAGCCCAATGATTCGTGCAACAGATPCGAAGCT 120
DB 215 TTCCCAAGCCCTTGTGATTTTGTGTAATAAACCCCATGATCTTCATCAGATAGAGCT 274
QY 121 TTCAAGAACTTCGGGTACAGAGTGGTTTGGCTATCAACTATCGCCAGAGGTAATG 180
DB 275 CTTAGGCAGTTGGAGTTGATGAATGGTTTGGCCATCAATTTATCAGCCAGAGGTGAT 334
QY 181 ATTAATTTCTGAAGACATTTAGGATAAGCTTGGCATCACAAATACATGCTCCCAAGAG 240
DB 335 CTCGAATCTTCAGACATTTGAGACCAAGCTGGAATCAAAATCACTGCTCACAAGAG 394
QY 241 ACTGAGCCCTTAGGAACCCCTGGCCCTTGTGCTCTAGCAAGGACAAAGCTTGGGATGGA 300
DB 395 ACCGAGCCACTAGTACCCCTGCTCTGGCTCTAGCGAGAGACAAATTCCTTGATGA 454
QY 301 TCTGCCCCACCAATCTTCTCTCTCAACAGTATGATGATCAAGCGAATACCGATTTGCTGAA 360
DB 455 TCTGGAGCCCTTCTCTCTTAAACAGTATGATGATGATGATGATGATGATGATGATGAT 514
QY 361 CTCATCAAAATTCACAAGTGTATGTTGGTGTGAGGCAACAATATGCTACTAGGTGGAT 420
DB 515 ATGCTTGACATTCACAAATCTCAGCTGGGAGAGCCCTCCATTAATGATAAAGAGTGGAT 574
QY 421 GAACCATCAAAATACGGTGTGTTGGTTATGAGGAGGCAACTGGCAGGCTGCAAAAGCTTT 480
DB 575 GAACGTCGAATATGGAGTGTGTTATGGAAGAGACACTGGAGAGTGGAGAGTTT 634
QY 481 GTTGAGAGCCAAATAATATTTGTGGTAAAGATCAATGCTGGGATTTACTTACTGAAAC 540
DB 635 GTGGAAGCCAAACTGATGTAGTAAAGATCAAGATCAACGCTGGGATTTATCTTCTGAAC 694
QY 541 CCATGCTGCTTGACGGCATTCAGCTGAGGCCAACAATCAATTGAGAAAGAGGCTTCCCT 600
DB 695 CCATCTGTTCTTGAATGAATTCAGCTAAGCCGATTCAGCTCAATCGAAAGAGACTTTCCCT 754
QY 601 CAAATTTGACGTGATCAAGACTATGCAATGGTCTTCCAGTTTGTGGATGGATGT 660
DB 755 AAGATTGCAGCAGCAAGGGCTCTATGCTATGGTCTACAGGGTTTGGATGGACAT 814
QY 661 GGTGAGCTTAGGACTACATTTACTGGCTTCGCTTCCTTATCTAGCTTCGATATAGGAAGAA 720
DB 815 GGGCAACCCCGTACTACATAACGGGTTTGAGACTCTACTTAGACTCCCTTAGGAAGAA 874
QY 721 TCAGCTGCAACCTAGCTACTGACACATGTTGTGGCAATGCTGCTGTCATGAGAC 780
DB 875 TCTCTGCCAATTAACCAAGTGGGCCACACATAGTTGGGAATGTTCTTGTGACGAAAC 934
QY 781 GCCAAGATTGGAGAAGTTGCTGATTTGCTGATGTCGCAATTTGACCTGGGCTGGTGT 840
DB 935 GCTACAATTTGGGAAGGATGTTGATTTGATGGACACAGACGTTGCCATTTGGTCCAGGCTCAT 994
QY 841 GTGGAGGAGCGGTGAGGCTTTCCCGCTGCACTGTGATGCGCGCGCTGCTATCAAGAAG 900
DB 995 GTTGAGTCAGGAGTCAGACTCTCCCGATGACGGGTCATGCTGGAGTCCGCTCAGAG 1054
QY 901 CATGCTTCGATTCGAGCAGTATCATCGGGTGGCACTCAACGCTTGGTCAATGGGACAGG 960
DB 1035 CATGCTGTATCTCAGCAGTATCATCGGGTGGCACTCAACGCTTGGTCAATGGGACAGG 1114
QY 961 ATAGAAATATGACTATCTCGGGGAGAGTGTCTATGTTGTGATGAGGTGTACAGCAAT 1020
DB 1115 ATCAGACACATGACATCTCTCGGAGGATGTTCTATGTCAGGAGATGATGATGATGAT 1174
QY 1021 GCGGTGTGTTCTCCACATCAAGAGATCAAGTCAAGCATTTCTGAAGCCTTGAGATCGTC 1080
DB 1175 GAGAGCTGTTTTCACACAGAGAGATCAAAATCAACATCTTGAAGCCAGAGATG 1234
QY 1081 ATGTGA 1086
DB 1235 ATGTGA 1240
```

```
RESULT 5
AAAX77618
ID AAX77618 standard; cDNA to mRNA; 1478 BP.
XX
AC AAX77618;
XX
DT 13-AUG-1999 (first entry)
XX
DE Potato GDP-mannose pyrophosphorylase cDNA.
XX
KW GDP-mannose pyrophosphorylase; transgenic plant; potato; plant cell;
KW GMPase; aerial part; ripening; ss.
XX
OS Solanum tuberosum.
XX
PH Key Location/Qualifiers
FT CDS 136..1221
FT /tag= a
FT /product= "GDP-mannose pyrophosphorylase"
XX
PN DE19741375-A1.
XX
PD 01-APR-1999.
XX
PF 19-SEP-1997; 97DE-1041375.
XX
PR 19-SEP-1997; 97DE-1041375.
XX
PL (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX
PI Keller R. Kossmann J;
XX
WPI: 1999-Z15846/19.
P-PSDB: AAY08879.
XX
DNA encoding potato GDP-mannose pyrophosphorylase protein - and
corresponding protein, antibody, transgenic plants with decreased or
increased GDP-mannose pyrophosphorylase activity, etc.
XX
Claim 2b: Page 11-13: 22pp; German.
XX
This invention describes a method for the production of a novel
transgenic plant containing plant cells that have been genetically
modified to decrease their endogenous GDP-mannose pyrophosphorylase
(GMPase) activity. The invention also describes a method for producing
transgenic plants whose aerial parts mature, ripen and die earlier and
more rapidly than those of corresponding wild-type plants, comprising
genetically modifying a plant cell by introducing a foreign DNA molecule
to decrease the cell's endogenous GMPase activity, regenerating a plant
from the cell, and optionally producing further plants from this plant.
XX
Sequence 1478 BP; 395 A; 279 C; 334 G; 470 T; 0 other;
Query Match 59.9%; Score 650.8; DB 20; Length 1478;
Best Local Similarity 75.0%; Pred. NO. 1.7e-197;
Matches 814; Conservative 0; Mismatches 272; Indels 0; Gaps 0:
QY 1 ATGAAGGCCCTCATTTCTTGTGCGGGGTTTCGGAACCCGCCCTTCGGCTTTTGACTCTGAGC 60
DB 136 ATGAAGGCCATTTATCTGTTGGAGGTTTCGGTACTCGGCTCAGGCCACTCACCCCTCAGT 195
QY 61 TTCCCCAAACCCCTCTGGGATTTTGCAGAACCCCATGATTTGCAACAGCCATGATTTGCAACAGAGCT 120
DB 196 GTCCCAAGCCAGCTCTCGAAATTTGCTAAGCAAAACCAATGATTTTGCATCAGATGAGGCT 255
QY 121 TTCAAGAACTTCGGGTACAGAGTGGTTTGGCTATCAACTATCGCCAGAGGTAATG 180
DB 256 CTCAGGCTGTGGAGTACCGAAGTTGTACTGGCTATTAACTACCAACCTGAGGTGATG 315
QY 181 ATTAATTTCTGAAGACATTTAGGATAAGCTTGGCATCACAAATACATGCTCCCAAGAG 240
DB 316 CTCGAATCTTCAGACATTTGAGACCAAGCTGGAATCAACATCTTGAAGCCAGAGATG 375
```


Qy	241	ACTGAGCCCTTAGGAACCGCTGGCCCTCTTGCTCTAGCAAGGACAAGCTTGGGGATGGA	300
Db	376	ACTGAACCACTTGGCACTGCTGGTCCCTTGCTTGGCTAGAGATAAGCTGATAGATGAC	435
Qy	301	TCTGCCCAAGCCATCTTTGTCTCAACAGTGTATGTCATAAGCGAATACCACTTTCGTGA	360
Db	436	TCTGGTGAAACATTTTTTGTCTTAAACAGTGTATGTCAGTGAATATCCATTCGAAGAG	495
Qy	361	CTCATCAAAATTTACAAGTGCATGGTGTGAGGCAACAATATATGCTACTAAGTGGAT	420
Db	496	ATGATTCATTCACAAATCCCATGGTGTGAGGCTTCTTGTATGGTGACCAAGTGGAT	555
Qy	421	GAACCATCAAAATACGGTGTGTGTATGAGAGGCAACTGGCAGGCTGGAAAGGTTT	480
Db	556	GAGCCTTCTAAATATGGTGTGTGTCTATGGAAGAATCCACTGGGCAAGTAGAGATTT	615
Qy	481	GTTGAGAGCCAAAATATTTGTGGGTACACAGATCAATGCTGGGATTTACTTACTGAAC	540
Db	616	GTGGAGAAGCCAAAGTTATTTGTGGCAACAAGATCAATGCTGGATTTTACCTGCTGAAC	675
Qy	541	CCATCTGTCTTGACCGCATTGAGCTGAGGCCAACATCAATGAGAAAGAGTCTTCCCT	600
Db	676	CCTTCTGTCTAGACAGAAATTCATACGCCCAACATCAATTGAGAAAGAGTTTTCCA	735
Qy	601	CAAAATGCACTGATCAACAGCTCTATGCAATGGTCCCTTCCAGGTTTTTGGATGGATGT	660
Db	736	AAAATGCAAGAGAAAGAACTGATGCAATGGTGTACCTTGGATTTGGATGGACATT	795
Qy	661	GGTCAGCCTAGGACTACATACTGGCTGGCTCTTTATCTAGACTCGATTAGGAAGAA	720
Db	796	GGCCAACCAAGAGATTACATTAAGTCTGAGCTCTAGACTCTATCTGGATTCTTTAAAGAAACAC	855
Qy	721	TCAGCTGCAAGCTAGCTACTGGACACATGTTGTTGGCAATGTGCTGGTGCATGAGAC	780
Db	856	TCTTCACCTAAATGGCTTCAGGAGCCACACATTTGCGGAAATGTCTAGTGGATGAATCT	915
Qy	781	GCCAAAGTTGGAGAAGTTGTCTGATGCTCCTGATGTGCCATTTGGACCTGGGTGTGT	840
Db	916	GCCAAAGTTGGAGAAGTTGTGTATAGGACCAGATGTTGCAATTTGGTCTGGTGTGTG	975
Qy	841	GTGGAGAGCGCGTGAGGCTTTCCCGCTGCATGTCAATCGCGCGGTGCGTATCAAGAA	900
Db	976	ATTGAGTCTGGAGTTAGACTCTCCCGTTGCATGTGTATGCGAGGAGTCCGCATCAAGAA	1035
Qy	901	CATGTTTGCATCTCAACAGCAATTAATCGCTGGCACTCAACTTGGTCAATGGGCACGG	960
Db	1036	CATGATGCATCTCAGTAGCATATGCGCTGGCACTCTACTGTGGCAATGGGCTCGT	1095
Qy	961	ATAGAGAATATGACTATCTCGGGGAGGATGTTCTGTTGATGAGGTGTACAGCAAT	1020
Db	1096	GTTGAGAACATGACCATTTCTCGGGGAAGATGCCATGTTTGTGATGAAATTTACAGCAAT	1155
Qy	1021	GCGGTGTGTTTCTCCCAATAAAGAGATCAAGTCAAGCATCTCAAGCCTGAGATCGTC	1080
Db	1156	GGAGGTGTAGTTTTCGCCCAAGAGAGATCAATCCAGTATATTGAAACCTGAAATCGTG	1215
Qy	1081	ATGTGA 1086	
Db	1216	ATGTGA 1221	
RESULT 6			
AAC39741			
ID	AAC39741 standard; DNA; 556 BP.		
XX			
AC	AAC39741;		
XX			
DT	17-OCT-2000 (first entry)		
XX			
DE	Zea mays DNA fragment SEQ ID NO: 25732.		
XX			
KW	Hybridisation assay; genetic mapping; gene expression control;		

KW	protein identification; signal transduction pathway; metabolic;
KW	pathway; promoter; termination sequence; corn; ss.
XX	
OS	Zea mays subsp. mays.
XX	
PN	EP1033405-A2.
XX	
PD	06-SEP-2000.
XX	
PF	25-FEB-2000; 2000EP-0301439.
XX	
PR	25-FEB-1999; 99US-0121825.
PR	05-MAR-1999; 99US-0123380.
PR	09-MAR-1999; 99US-0123548.
PR	23-MAR-1999; 99US-0125788.
PR	25-MAR-1999; 99US-0126264.
PR	29-MAR-1999; 99US-0126785.
PR	01-APR-1999; 99US-0127462.
PR	06-APR-1999; 99US-0128234.
PR	08-APR-1999; 99US-0128714.
PR	16-APR-1999; 99US-0129845.
PR	19-APR-1999; 99US-0130077.
PR	21-APR-1999; 99US-0130449.
PR	23-APR-1999; 99US-0130510.
PR	23-APR-1999; 99US-0130891.
PR	28-APR-1999; 99US-0131449.
PR	30-APR-1999; 99US-0132048.
PR	30-APR-1999; 99US-0132407.
PR	04-MAY-1999; 99US-0132484.
PR	05-MAY-1999; 99US-0132485.
PR	06-MAY-1999; 99US-0132486.
PR	06-MAY-1999; 99US-0132487.
PR	07-MAY-1999; 99US-0132863.
PR	11-MAY-1999; 99US-0134256.
PR	14-MAY-1999; 99US-0134218.
PR	14-MAY-1999; 99US-0134219.
PR	14-MAY-1999; 99US-0134221.
PR	14-MAY-1999; 99US-0134370.
PR	18-MAY-1999; 99US-0134768.
PR	19-MAY-1999; 99US-0134941.
PR	20-MAY-1999; 99US-0135124.
PR	21-MAY-1999; 99US-0135353.
PR	24-MAY-1999; 99US-0135629.
PR	25-MAY-1999; 99US-0136021.
PR	27-MAY-1999; 99US-0136392.
PR	28-MAY-1999; 99US-0136782.
PR	01-JUN-1999; 99US-0137222.
PR	03-JUN-1999; 99US-0137528.
PR	04-JUN-1999; 99US-0137502.
PR	07-JUN-1999; 99US-0137724.
PR	08-JUN-1999; 99US-0138094.
PR	10-JUN-1999; 99US-0138540.
PR	10-JUN-1999; 99US-0138847.
PR	14-JUN-1999; 99US-0139119.
PR	16-JUN-1999; 99US-0139452.
PR	16-JUN-1999; 99US-0139453.
PR	17-JUN-1999; 99US-0139492.
PR	18-JUN-1999; 99US-0139454.
PR	18-JUN-1999; 99US-0139455.
PR	18-JUN-1999; 99US-0139456.
PR	18-JUN-1999; 99US-0139457.
PR	18-JUN-1999; 99US-0139458.
PR	18-JUN-1999; 99US-0139459.
PR	18-JUN-1999; 99US-0139460.
PR	18-JUN-1999; 99US-0139461.
PR	18-JUN-1999; 99US-0139462.
PR	18-JUN-1999; 99US-0139463.
PR	18-JUN-1999; 99US-0139750.
PR	18-JUN-1999; 99US-0139763.
PR	21-JUN-1999; 99US-0139817.
PR	22-JUN-1999; 99US-0139899.
PR	23-JUN-1999; 99US-0140353.
PR	23-JUN-1999; 99US-0140354.

PR	24-JUN-1999;	99US-0140695.	PR	07-SEP-1999;	99US-0152363.
PR	28-JUN-1999;	99US-0140823.	PR	10-SEP-1999;	99US-0153070.
PR	29-JUN-1999;	99US-0140991.	PR	13-SEP-1999;	99US-0153758.
PR	30-JUN-1999;	99US-0141287.	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0141842.	PR	16-SEP-1999;	99US-0154039.
PR	01-JUL-1999;	99US-0142154.	PR	20-SEP-1999;	99US-0154779.
PR	02-JUL-1999;	99US-0142055.	PR	22-SEP-1999;	99US-0155139.
PR	06-JUL-1999;	99US-0142390.	PR	23-SEP-1999;	99US-0155406.
PR	08-JUL-1999;	99US-0142803.	PR	24-SEP-1999;	99US-0155659.
PR	09-JUL-1999;	99US-0142920.	PR	28-SEP-1999;	99US-0156458.
PR	12-JUL-1999;	99US-0143977.	PR	29-SEP-1999;	99US-0156596.
PR	13-JUL-1999;	99US-0143542.	PR	04-OCT-1999;	99US-0157117.
PR	14-JUL-1999;	99US-0143624.	PR	05-OCT-1999;	99US-0157753.
PR	15-JUL-1999;	99US-0144005.	PR	06-OCT-1999;	99US-0157865.
PR	16-JUL-1999;	99US-0144085.	PR	07-OCT-1999;	99US-0158029.
PR	19-JUL-1999;	99US-0144086.	PR	08-OCT-1999;	99US-0158232.
PR	19-JUL-1999;	99US-0144325.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144331.	PR	13-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-0144333.	PR	13-OCT-1999;	99US-0159295.
PR	19-JUL-1999;	99US-0144334.	PR	14-OCT-1999;	99US-0159329.
PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159330.
PR	20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159331.
PR	20-JUL-1999;	99US-0144632.	PR	14-OCT-1999;	99US-0159637.
PR	20-JUL-1999;	99US-0144884.	PR	14-OCT-1999;	99US-0159638.
PR	21-JUL-1999;	99US-0144814.	PR	14-OCT-1999;	99US-0159638.
PR	21-JUL-1999;	99US-0145086.	PR	18-OCT-1999;	99US-0159584.
PR	21-JUL-1999;	99US-0145088.	PR	21-OCT-1999;	99US-0160741.
PR	21-JUL-1999;	99US-0145085.	PR	21-OCT-1999;	99US-0160767.
PR	22-JUL-1999;	99US-0145087.	PR	21-OCT-1999;	99US-0160768.
PR	22-JUL-1999;	99US-0145089.	PR	21-OCT-1999;	99US-0160770.
PR	22-JUL-1999;	99US-0145192.	PR	21-OCT-1999;	99US-0160814.
PR	23-JUL-1999;	99US-0145145.	PR	21-OCT-1999;	99US-0160815.
PR	23-JUL-1999;	99US-0145218.	PR	22-OCT-1999;	99US-0160980.
PR	23-JUL-1999;	99US-0145224.	PR	22-OCT-1999;	99US-0160981.
PR	23-JUL-1999;	99US-0145224.	PR	22-OCT-1999;	99US-0160989.
PR	26-JUL-1999;	99US-0145276.	PR	22-OCT-1999;	99US-0160989.
PR	27-JUL-1999;	99US-0145913.	PR	25-OCT-1999;	99US-0161404.
PR	27-JUL-1999;	99US-0145918.	PR	25-OCT-1999;	99US-0161405.
PR	27-JUL-1999;	99US-0145919.	PR	25-OCT-1999;	99US-0161406.
PR	28-JUL-1999;	99US-0145951.	PR	26-OCT-1999;	99US-0161359.
PR	02-AUG-1999;	99US-0146386.	PR	26-OCT-1999;	99US-0161360.
PR	02-AUG-1999;	99US-0146388.	PR	26-OCT-1999;	99US-0161361.
PR	02-AUG-1999;	99US-0146389.	PR	28-OCT-1999;	99US-0161920.
PR	03-AUG-1999;	99US-0147038.	PR	28-OCT-1999;	99US-0161992.
PR	04-AUG-1999;	99US-0147204.	PR	28-OCT-1999;	99US-0161993.
PR	05-AUG-1999;	99US-0147302.	PR	29-OCT-1999;	99US-0162142.
PR	05-AUG-1999;	99US-0147192.			
PR	05-AUG-1999;	99US-0147260.			
PR	06-AUG-1999;	99US-0147303.			
PR	06-AUG-1999;	99US-0147416.			
PR	09-AUG-1999;	99US-0147493.			
PR	09-AUG-1999;	99US-0147935.			
PR	10-AUG-1999;	99US-0148171.			
PR	11-AUG-1999;	99US-0148319.			
PR	12-AUG-1999;	99US-0148341.			
PR	13-AUG-1999;	99US-0148565.			
PR	13-AUG-1999;	99US-0148684.			
PR	16-AUG-1999;	99US-0149368.			
PR	17-AUG-1999;	99US-0149175.			
PR	18-AUG-1999;	99US-0149426.			
PR	20-AUG-1999;	99US-0149722.			
PR	20-AUG-1999;	99US-0149723.			
PR	20-AUG-1999;	99US-0149920.			
PR	23-AUG-1999;	99US-0149902.			
PR	23-AUG-1999;	99US-0149930.			
PR	25-AUG-1999;	99US-0150566.			
PR	26-AUG-1999;	99US-0150804.			
PR	27-AUG-1999;	99US-0151065.			
PR	27-AUG-1999;	99US-0151066.			
PR	27-AUG-1999;	99US-0151080.			
PR	30-AUG-1999;	99US-0151303.			
PR	31-AUG-1999;	99US-0151438.			
PR	01-SEP-1999;	99US-0151930.			
Query Match 40.8%; Score 443; DB 21; Length 556;					
Best Local Similarity 98.9%; Pred. No. 3.6e-131;					
Matches 457; Conservative 0; Mismatches 1; Indels 4; Gaps 1;					
QY	1	ATGAAGCCCTCATCTTGTGCGGGGTTTCGGAACCGCCCTTCGGCCTTTCGACTCTGAGC	60		
Db	99	ATGAAGCCCTCATCTTGTGCGGGGTTTCGGAACCGCCCTTCGGCCTTTCGACTCTGAGC	158		
QY	61	TTCCCGAAACCCCTCGTGGATTTTGCACAAAGCCCATGATTTCTGCACAGATCGAAGCT	120		
Db	159	TTCCCGAAACCCCTCGTGGATTTTGCACAAAGCCCATGATTTCTGCACAGATCGAAGCT	218		
QY	121	TTGAACAAGCTTGGGGTTCACAGAGCTGGTTTGGCTATCAACTATPGCCCGCAGAGTAATG	180		
Db	219	TTGAAGAAGTTGGGGTTCACAGAGGTTGTTTGGCTATCAACTATPGCCCGCAGAGTAATG	278		
QY	161	ATTAAATTCCTGAAGGACTTGAGGATAAGCTTGGGATCAAAATACATGCTCCCAAGAG	240		
Db	279	ATTAAATTCCTGAAGGACTTGAGGATAAGCTTGGGATCAAAATACATGCTCCCAAGAG	338		
QY	241	ACTGAGCCCTTAGGAACCGCTGGCCCTCTTGTCTAGCAAGGGACAAAGCTTGGCGATGGA	300		
Db	339	ACTGAGCCCTTAGGAACCGCTGGCCCTCTTGTCTAGCAAGGGACAAAGCTTGGCGATGGA	398		
QY	301	TCTGGCCAGCCATCTTTTGTCTCAACAGATGATGATCAAGCGAATACCCATTGCTGAA	360		
Db	399	TCGTG---GCCATCTTTGTCTCAACAGATGATGATCAAGCGAATACCCATTGCTGAA	454		

QY 361 CTCATCAAAATTCACAGTGTCTATGGTGGTGAGGCAACAATATATGGTCACTAAGGTGGAT 420
|||||
Db 455 CTCATCAAAATTCACAGTGTCTATGGTGGTGAGGCAACAATATATGGTCACTAAGGTGGAT 514
|||||
QY 421 GAACCATCAAAATACGGTGTGTGGTTATGGAGGAGGCAACT 462
|||||
Db 515 GAACCATCAAAATACGGTGTGTGGTTATGGAGGAGGCAACT 556
|||||
RESULT 7
ABL20809
ID ABL20809 standard; DNA; 1358 BP.
XX
AC ABL20809;
XX
DT 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 13900.
DE
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
XX Drosophila melanogaster.
OS
XX WO200171042-A2.
PN
XX
XX
PD 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
PF
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE) PE CORP NY.
PA
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX
XX WPI; 2001-656860/75.
DR
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX Claim 1; SEQ ID NO 13900; 2lpp + Sequence Listing; English.
PS
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 1358 BP; 375 A; 317 C; 361 G; 305 T; 0 other;
SQ
Query Match 38.4%; Score 417.2; DB 23; Length 1358;
Best Local Similarity 62.8%; Pred. No. 1.1e-122;
Matches 682; Conservative 0; Mismatches 398; Indels 6; Gaps 2;
QY 1 ATGAAGGCCCTTCATCTGTTCGGGGTTTCGGAACCCGCCCTTCGGCCCTTGTACTCTGAGC 60
|||||
Db 279 ACGAGGGCGTGTATCTGGTCGGTGGTTATGGGACCCGCCCTTCGCCCTCAGC 338
|||||
QY 61 TTCGCGAAACCCCTCGTGGATTTTGCACAAAGCCCATGATCTGCACCATCAAGCT 120
|||||
Db 339 ACGCCCAAGCAGCGGTGGAGTTTGCACAAAGCCGATCTCTCCACCAACTGGAGGCA 398
|||||
QY 121 TTGAAGAAGTTGGGGTCACAGAGGTGGTTTTGGCTATCAACTATCGCCAGAGGTAATG 180
|||||

Db 399 CTCGTGATGCGGGATGTCTGTCAGGTTATTTTAGCCGTGAGCTATCGAGCGGACCAATG 458
|||||
QY 181 ATTAATTTCTTGAAGACCTTTGAGGATTAAGCTTGGGATCACAATTTACATGCTGCCAAGAG 240
|||||
Db 459 GAAAAGGAGCTTAAGTCGAGCAAGAACTGGCGTGAATTTGATCTCTACACGAG 518
|||||
QY 241 ACTGAGCCCTTAGGAACCGCTGCCCTCTTCTCTAGCAAGGACAAAGCTTGGCGATGGA 300
|||||
Db 519 ACGGAACCCCTGGGAACAGCTGGACCTCTGGCCCTAGCAAAAACGATTTTAGCAGCCAGC 578
|||||
QY 301 TCTGGCCAGCCATCTTTTCTCTCAACAGTGTATAGCAAGCAATPACCCATTTGCTGAA 360
|||||
Db 579 TCAG---AGCCATTTTTCGTGCTCAATTCGGAGCTTATATGCGATTTTCCATTTAAACAA 635
|||||
QY 361 CTCATCAAAATTCACAAGTGTCTATGGTGGAGCAACAATTTATGGTCACTAAGTGGAT 420
|||||
Db 636 CTAGTGCATTTCCATTTGATACGGAAGAAGAGGCAACAATTTGTCACAAGAGTGCAG 695
|||||
QY 421 GAACCATCAAAATACGGTGTGTGGTTATGGAGGAGCAACTGGCAGGGTGGAAAGTTT 480
|||||
Db 696 GAACCATCAAAATACGGAGTTGTG---CTGTACGATGAGAAGCGCTGTATAAAAAACTTT 752
|||||
QY 481 GTTGAGAAGCCAAAATAATTTGTGGGTAAACAAGATCAATGCTGGGATTTACTTACTGAAC 540
|||||
Db 753 ATTGAAAAGCCACAAGAGTTCTGTAGCAATAAGATCAATGCCGCAATTTACATCTTTAAT 812
|||||
QY 541 CCATCTGTCTTGACCGCATTTGAGCTGAGGCCCAACATCAATTTGAGAAAGAGTCTTCCCT 600
|||||
Db 813 CCGTCGTGCTTGACCGGATTCGAAAGTTAAGCCACATCAATAGAGAGGAGTGTCCCT 872
|||||
QY 601 CAAATTCAGCTGATCAACAGCTCTATGCAATGGTCTTCCAGGTCTTTCGATGGATGTT 660
|||||
Db 873 GAAATGACCGCAACAGGAGTTGTATGCAATGGAATTAAGTCTGGATGGACATC 932
|||||
QY 661 GGTACGCTTAGGAGTACATTTACTTGTGCTGCTTTATCTAGACATGATAGAGAGAA 720
|||||
Db 933 GGACAGCCAAAAGACTTCCCTAACCGCATGTGTCTGTATCTAAGCTCGCTGCCAGAG 992
|||||
QY 721 TCAGCTGCCAAGCTAGCTACTGAGCAGCATGTTTGGCAATGCTGTGGTCATGAGAGC 780
|||||
Db 993 CAATCCCCCAAGCTGTACACAGGCCCTGGAGTGGTGGCAACGTTGTGGGATPCCACG 1052
|||||
QY 781 GCCAAGATTGGAGAAGTTGTCTGATTGTCCTGATGTGCCATTTGGACCTGGGTGTGTT 840
|||||
Db 1053 GCCAAATCGCGAGGTTGTCCATAGGCTTAATGTAACCATTTGACCGGAGCTGCTC 1112
|||||
QY 841 GTGGAGACGGGTGAGGCTTTCCCGCTGCATGTATGCGCGCGTGCCTATCAAGAG 900
|||||
Db 1113 ATCGAGGATGGCGTGTGCATTAAGCGCTCGACCATCTCAAAGCGCCATCGTTCGCTCA 1172
|||||
QY 901 CATGCTTGATCTCAACAGCATTTATCGGTGGCAGCTCAACTGTTGTCATGGGACGG 960
|||||
Db 1173 CACTCATGCTGGACTCCTGCATTTGCGTTGCGTGTCTACAGTTGGCCGTTGGTTCGC 1232
|||||
QY 961 ATAGACAATATGACTACTCTCGGGGAGGATGTTTCATGTGTGATGAGGTACAGCAAT 1020
|||||
Db 1233 ATCGAAGGATCACCGTGTAGCGGAGGATGTAATCGTGAAGGATGAGCTCTACATTAAT 1292
|||||
QY 1021 GCGGTGTTTCTCCACATAAAGAGATCAAGTCAAGCATTTCTGAAGCCCTGAGATCGTC 1080
|||||
Db 1293 GGAGGCCAGGTCTGCCCCATAAAGCATTTGCGGCCAGTGTCCAGAACCCGAGATCAT 1352
|||||
QY 1081 ATGTGA 1086
|||||
Db 1353 ATGTGA 1358
|||||

RESULT 8
ABL02105
ID ABL02105 standard; cdna; 1476 BP.
XX
AC ABL02105;

XX 26-MAR-2002 (first entry)
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 797.
DE Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene: ss.
KW Drosophila melanogaster.
XX WO200171042-A2.
PN 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US09231.
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX (PEKE) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-658660/75.
DR P-PSDB; ABB58002.
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX Claim 1: SEQ ID NO 797; 21pp + Sequence Listing; English.
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (AB557737-AB572072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 1476 BP; 428 A; 335 C; 380 G; 333 T; 0 other;
2;
Query Match 38.4%; Score 417.2; DB 23; Length 1476;
Best Local Similarity 62.8%; Pred. No. 1.2e-122;
Matches 682; Conservative 0; Mismatches 396; Indels 6; Gaps 2;
Qy 1 ATGAAGGCCCTCATCTCTGTCGGGGTTCGGAACCCGCTTCGGCCTTTGACTCTGAGC 60
Db 397 ACGAGGGCGTTGATTTCTGTCGGTGGTTATGCGGACCGCGCTCGACCCCTGACCTCAGC 456
Qy 61 TTCGCGAAACCCCTCGGATTTGCAACAAGCCCATGATTTCTGCACCAGATCGAAGCT 120
Db 457 ACCGCCAAGCCACTCGTCAGATTTGCCAATAAGCCGATTTCTCTCCACCAACTGGAGCA 516
Qy 121 TTGAAGAAGATTTGGGGTTCACAGAGTGGTTTGGCTATCAACTATCGCCCGAGAGTAAATG 180
Db 517 CTCGTCGATCGGGATGTCGTAGGTATTTTAGCGTCAGCTATCGAGCGGAGCAATG 576
Qy 181 ATTAAATTTCTGAAGCATTTGAGATTAAGCTTGGCATCACAAATPACATGCTCCCAAGAG 240
Db 577 GAAAGAGAGCTTAAATGTCGAAGCCAAAGAACTGGGCGTGGAAATGATCTTCTCACACGAG 636
Qy 241 ACTGAGCCCTTAGGAACCGCTGGCCCTCTGCTCTAGGAAGGACAAAGCTTTCGGGATGGA 300
Db 637 ACGGAACCCCTGGGAACAGCTGGAGCTCTGGCCCTTAGCAAAACCAATTTAGCAGCCAGC 696
Qy 301 TCTGCCACGCAATTTTGTCTCTCAACAGTGTATGTCATAAGCGAATACCCATTTGCTGAA 360
Db 697 TCAG---AGCCATTTTCTGCTCAATTCGGACGTTATATGCGATTTTCCATTTTAAACAA 753

QY 361 CTCATCAAAATTTACAAAGTGTCTATGGTGGTGAAGGCAACAATATGGTCACTAAGGTGAT 420
Db 754 CTAGTGCATTTCCATTTCTAATCAGCGAAAAGAGCGCAATTTCTGTCCACAAAGTCGAA 813
QY 421 GAACCATCAAAATACGGTGTGTGGTTATGGAGGAGGCAACTGCGAGGTGGAAAGGTTT 480
Db 814 GAACCATCAAAATACGGAGTTGTG---CTGTACGATGAGAAGCGCTGTATAAAACATTT 870
QY 481 GTTGACAAGCAAAAATATTTTGTGGTAAACAAGATCAATGCTGGATTTACTTACTGAAC 540
Db 871 ATTGAAGCCCAAGAGTTCGTTAGCAATAGATCAATGCCGCAATTTACATCTTTAAT 930
QY 541 CCATCTGCTCTTGACCGCATTTAGCTGAGCGCAACATCAATGAGAAGAGGCTTCCCT 600
Db 931 CCCTCCGCTGTGACCGCATCGAAGTTAAGCCCAACATAGACAGAGGAGTGTCCCT 990
QY 601 CAATTCAGCTGATCAACAGCTCTATGCAATGCTCTTCCAGGTTTTTGGATGATGAT 660
Db 991 GAAATGACGAGCAACAGAGGTTGTATGCAATGATTAAGTCTGCTTAAAGCTCGCTGCGCAGAAG 1050
QY 661 GCTCAGCTAGGCACTACATTTACTGCTTCCGCTTTTATCTAGACTCGATTAGAGAGAA 720
Db 1051 GGACAGCCAAAGACTTCTTAACCGCATGTGCTGTATCTAAGCTCGCTGCGCAGAAG 1110
QY 721 TCAGCTGCCAAGCTAGCTACTGAGGACACATGTTTGGCAATGCTGGTGCATGAGAGC 780
Db 1111 CAATCCCCCAAGCTGTACACAGGCCCTGAGTGTGGCAACGTTGGTGGCATCCACG 1170
QY 781 GCCAAGATTTGGAGAAGTTTCTGATTTGTCCTGATGTGCGCATTTGACCTGGGTGTGTT 840
Db 1171 GCCAAATCGCGAGGGTTCTGCATAGGCGCTAATGTAACCATTTGGACCGGACGTGTC 1230
QY 841 GTGAGAGCGCTGAGGCTTTCCCGCTGCATGCTATGCGCGCGCTGCTGATATCAAGAG 900
Db 1231 ATCGAGGATGGCTGTGCAATTAAGCGCTGACCATCTCTCAAGCGCCATGCTTCCGCTCA 1290
QY 901 CATGCTTGCATCTCAACACATATCGGCTGGCACTCAACTGCTGCTCAATGGCAGCG 960
Db 1291 CATCATGCTGACATCTCTGATTTGCGTTGGCGTTCTACAGTTGGCCGTTGGTGTGCG 1350
QY 961 ATAGAGAATATGACTATCTCTGGGGAGGATGTTTCATGTGTGATGAGGTGATAGCAAT 1020
Db 1351 ATCGAAGGCATCACCGTGTGCTAGGCGAGGATGTAATCTGGAAGGATGAGTCTACATTAAT 1410
QY 1021 GCGGTTGTTCTTCCACATAAGAGATCAAGTCAAGCATTTCTGAGAGCTGAGATCGTC 1080
Db 1411 GGAGGCGAGCTTGGCCCAATAAAGCATTTGCGGCCAGTGTCCCAAGAACCGCAGATCAT 1470
QY 1081 ATCTCA 1086
Db 1471 ATGTGA 1476
RESULT 9
AAC75966
ID AAC75966 standard; cDNA; 1512 BP.
XX AAC75966;
XX 08-FEB-2001 (first entry)
XX Human ORF1521 polynucleotide sequence SEQ ID NO:3041.
DE Human; open reading frame: ORF; detection: cytostatic; hepatotropic;
XX vulvar; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antifibrinolytic;
KW hypotensive; dermatological; immunosuppressive; antitumor; antithyroid;
KW antiviral; antibacterial; antifungal; antirheumatic; antidiabetic;
KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;

XX Transferase; TRNSES-2; human; antitumour; cancer; inflammation;
KW gastrointestinal disorder; developmental disorder;
KW genetic disorder; neurological disorder; reproductive disorder;
KW smooth muscle disorder; immunological disorder; gene therapy;
KW diagnosis; mannose-1-phosphate guanyltransferase; ss.
XX Homo sapiens.
OS
FH Key Location/Qualifiers
FT CDS 213..1295
FT /*tag- a
XX
PN WO200014251-A2.
XX
PD 16-MAR-2000.
XX
PF 09-SEP-1999; 99WO-US20989.
XX
PR 10-SEP-1998; 98US-0150657.
PR 04-NOV-1998; 98US-0186779.
PR 11-MAY-1999; 99US-0133642.
XX (INCY-) INCYTE PHARM INC.
XX Tang YT, Corley NC, Guegler KJ, Baughn MR, Lal P, Yue H;
PI Hillman JL, Azimzai Y;
XX
DR WPI: 2000-256996/22.
DR P-PSDB; AAY79210.
XX
XX Human transferase proteins useful for preventing, diagnosing and
PT treating cancers and developmental, gastrointestinal, genetic,
PT immunological, neurological, reproductive and smooth muscle disorders -
XX
PS Claim 9; Page 97-98; 113pp: English.
XX
CC The present sequence is that of cDNA clone 2682663B1 encoding
CC human transferase TRNSES-2 (see AAY79210), 1 of 15 claimed human
CC transferase proteins of the invention (see AAY79209-23). The clone
CC was isolated from ileum tissue cDNA library SINDUCT01. TRNSES-2
CC is expressed in reproductive and gastrointestinal tissues,
CC especially those associated with cancer and inflammation. It
CC shows homology to mannose-1-phosphate guanyltransferases. A
CC polynucleotide comprising nucleotides 377-406 of the present
CC sequence can be used as a DNA probe. The new human transferase
CC proteins and the polynucleotides encoding them can be used in the
CC diagnosis, prevention and treatment (including gene therapy and
CC antisense therapy) of cancer, developmental disorders,
CC gastrointestinal disorders, genetic disorders, immunological
CC disorders, neurological disorders, reproductive disorders, and
CC smooth muscle disorders.
XX
SQ Sequence 1559 BP; 304 A; 429 C; 519 G; 307 T; 0 other;
XX
Query Match 36.8%; Score 399.6; Db 21; Length 1559;
Best Local Similarity 61.1%; Pred. No. 5.3e-117;
Matches 664; Conservative 0; Mismatches 419; Indels 3; Gaps 1;
OY 1 ATGAAGGCCCTCATCTTGTCTGGGGGTTTCGGAAACCGGCTTGGGCTTTGACTCTGAGC 60
DB 213 ATGAAGGCACTGATCTAGTGGGGGCTATGGGACCGCGCTACGGCGGCTGAGCGCTGAGC 272
OY 61 TTTCGCAACCCCTCTGCAATTTTCGAACACAGCCCATCATTTTCACACAGATCGAAGCT 120
DB 273 ACCCGGAGGCCACTGCTGGAGCTTCTGCAATAAGCCCATCTTGTCTGCAACCAAGTGGAGCG 332
OY 121 TTGAAGAAGTTGGGCTCACACAGCTGCTTTGGCTATGAAGTATGCGCCACAGCTAATG 180
DB 333 CTAGCCGGGCGAGCGGTGGACCCAGCTGATCTTGGCGGTGAGCTACATGTCGCGAGGTGCTG 392
OY 181 ATTAATTTCTTAAGGACCTTTCAGGATAAGCTTGGCATCACAATTAATCATGCTCCCAAGAG 240
XX

DB 393 CACAAGCAATGAAGCCACAGGACACAGCGCTGCCAATCCCAATCTCCATCTCCCATCAA 452
OY 241 ACTGAGCCCTTAGGAACCGCTGGCCCTCTTGTCTTAGCAAGGGACAGCTTTCGGGATGGA 300
DB 453 GAGGAGCCTTTGGGACAGCTGGGCCCTTGGCGCTGGCCGTGACCTACTCTC---TGAG 509
OY 301 TCTGGCCAGCCATTCTTTGCTCTCAACAGCTGATCTGCGATTTCCCTTCCCAAGCC 360
DB 510 ACTGCAGACCTTTCTTCGTCTCAACAGTGACGTGATCTGCGATTTCCCTTCCCAAGCC 569
OY 361 CTCATCAAAATTCACAAGTGTGATGGTGAGGCAACAATATATGGTCACTCAAGGTGAT 420
DB 570 ATGGTCAGTTCCACCGGACCATGGCCAGGAGGCTCCATCCCTGGTGACCAAGGTGGAG 629
OY 421 GAACCATCAAAATACGGTGTGTGGTTATGGAGGAGGCAACTGGCAGGGTGTGAAGGTTT 480
DB 630 GAACCTCCCAAGTACGGTGTGGTGTGTGAGGCTGACAGCCGCTTCCACGGGTTT 689
OY 481 GTTGACAAGCCAAAATAATTTGTGGGTPAACAAAGATCAATATGCTGGGATTTACTTGAAC 540
DB 690 GTGGAGAAGCCACAGGTGTTTGTCTCAATAAGATCAACGAGGATGTACATCTCAGC 749
OY 541 CCATCTGCTTTCACCGCATTTGAGCTGAGGCCCAACATCAATTGAGAAAGAGGCTCTCCCT 600
DB 750 CTTGCAAGTGTCTGGCGGCATCCAGCTGCAGCTACGCTTACGTTGAGAGGAGGTCTTCCCC 809
OY 601 CAAATTCAGCTGATCAACAGCTCTATGCAATGGTCTCTCCAGGTTTTCAGATGATGTT 660
DB 810 ATTATGCCAAGAGAGGGGAGCTATATGCCATGAGTACAGGCTTCTGGATGGACATT 869
OY 661 GGTGAGCCTAGGACTACATTACTTGGCTTCGCTCTTTATCTAGACTCGATTAGGAAGAA 720
DB 870 GGGCAGCCCAAGGACTTCTCTCACTGGCATGTGCTCTTCTGTCAGTCACTGAGGAGAA 929
OY 721 TCAGTCCCAAGCTAGTACTGAGGACACATGTTGCGCAATCTGCTGGTCTGCTCAGAGC 780
DB 930 CAGCCTCAGCGGCTGTGCTCAGGCCCTTGGCATTTGTGGGCAACGTCGTGGTGGACCAAGT 989
OY 781 GCCAAGATTGGAAAGGTTGCTGATTGTCCTGATGTGCGCATTTGGACCTGGGTGTGTT 840
DB 990 CCGCGCATCCGACAGAACTGACAGCAATTTGGCCGCCAATGTCAGCTGCGGCTGGTG 1049
OY 841 GTGGAGGACGGCTGAGGCTTTCCCGCTGCACTGTCTCATGCGCGCGTGGGTATCAAGAAG 900
DB 1050 GTCGAAGATGCTGTGTATCCGCGGTGTCAGCGGTGTCGGGATGCCGATCCGCTTCC 1109
OY 901 CATGCTGCAATGCAACAGCAATTTCCCGCTGCACTGTCTCATGCGCGCGTGGGTATCAAGAAG 960
DB 1110 CATTCCTGGCTTGGTCTGCTGATTTGGGCTGGCGCTGCGCGGTGAGTGGGTACGC 1169
OY 961 ATAGAGAAATGACTATCTCCGGGAGGATGTTTCATCTGTGATGAGTGTACAGCAAT 1020
DB 1170 ATGGAAGACCTGACATGCTGGTGGGAGAGCTGCTGATGATGATGATGATGATGATGAT 1229
OY 1021 GCGGCTGTTTCTTCTCCACATATAAGAGATCAAGTCAAGGATTCCTGAAGCCCTGAGATGTC 1080
DB 1230 GGAGCCAGCGTGTGCGCCACAGCTCTATTGGGAGTCTATTGGGAGTCTGCTATCATC 1289
OY 1081 ATGTGA 1086
DB 1290 ATGTGA 1295
RESULT 11
ABA01227
ID ABA01227 standard; DNA; 1092 BP.
XX
XX ABA01227;
XX
XX 25-JAN-2002 (first entry)
XX
XX NDP-hexose pyrophosphorylase homolog coding sequence.
XX

KW Geminivirus; plant; viral infection; transgenic plant;
XX tomato yellow leaf curl virus; ds.
XX Schizosaccharomyces pombe.
OS FR2806095-A1.
XX 14-SEP-2001.
PD 10-MAR-2000; 2000FR-0003140.
XX 10-MAR-2000; 2000FR-0003140.
XX 10-MAR-2000; 2000FR-0003140.
XX (GENT-) GENTECH SARL.
XX Bejarano ER, Castillo GA, Colinet D, Donoso CI, Iniesta JR;
PI Grevesse C, Hericourt F;
XX WPI; 2001-628275/73.
DR P-PSDB; AAM52357.
XX New polynucleotides for producing transgenic plants resistant to
XX geminivirus infection comprising polynucleotides encoding proteins
XX which interact with at least one of the products of the geminivirus
XX genome -
XX
XX Claim 5; Pages 36-37; 106pp; French.
XX
XX The present invention relates to coding sequences encoding proteins which
XX interact with at least one of the six products of the geminivirus genome
XX necessary for infection of a plant by the virus. The present sequence is
XX the coding sequence for one such protein. This sequence is useful for
XX producing transgenic plants resistant to geminivirus infection,
XX particularly tomato yellow leaf curl virus.
XX
XX Sequence 1092 BP; 253 A; 230 C; 250 G; 359 T; 0 other;
XX
XX Query Match 33.0%; Score 358; DB 22; Length 1092;
XX Best Local Similarity 60.2%; Pred. No. 9, 1e-104;
XX Matches 649; Conservative 0; Mismatches 420; Indels 9; Gaps 3;
Qy 1 ATGAAGCCCTCATCTCTTGGGGGTTTCGGAACCCGCTTCGGCTTTGACTCTGAGC 60
Db 1 ATGAAGCCCTCATCTCTTGGGGGTTTCGGAACCCGCTTCGGCTTTGACTCTTAACT 60
Qy 61 TTCGGAACCCCTCGTGATTTTCGAACCAAGCCGATGATCTGCACAGATCGAAGCT 120
Db 61 TTCGGAACCCCTCGTGATTTTCGAACCAAGCCGATGATCTGCACAGATCGAAGCT 120
Qy 121 TTGAAGAAGTTGGGGTACAGAGCTGTTTGGCTATCAACTATCGCCAGAGGTAATG 180
Db 121 TTGGCAGCTGCTGGGTACAGGATATCGTCTTGGCTGTTAATACCGTCCCGAAATCATG 180
Qy 181 ATTAATTTCTGAAGGACTTTGAGGATAAGCTTGGCATCACAATTAACATGCTCCCAAGAG 240
Db 181 GTAGAAGCTTTGAAAATACCAAAAGGAGTATACGTTAATAATTAATCTTCCGTTGAG 240
Qy 241 ACTGAGCCCTTAGGAACCCGCTGGCCCTCTGCTCTAGCAAGGACAAAGCTTCGGGATGA 300
Db 241 AATGAGCCCTTTGGGAACCCGCTGGACCTTTGGCTCTGCTCGTGACATTTCTTGC -TAA 297
Qy 301 TCTGCCAGCCATTTCTTGTCTCAACAGTATGATCATAGGAATACCCATTTGCTGAA 360
Db 298 GACCACCTCTCTCTTCTGCTCTGTAATAGTATGATTTTGTGAGTATCTTCCGCGGAT 357
Qy 361 CTCATCAATTTCAAGTGTGATGCTGAGGCAACAAATATGCTACTAAGGTGAT 420
Db 358 TTGGCAGCTTTCCATAAGGCTCATGTCGCGAGGGTACTATGTTTACCAGGTTGAG 417
Qy 421 GAACCAATAAATACGGTGTGTTGTTATGGA -GGAGGCAACTGGCAGGTTGGAAGG 477
Db 418 GAGCCTTCTAAGTAGTGGTGTGCTGCTTATTCATATCCCAACTCAGATCTTGTATGAGCGC 477

Qy 478 TTTGTTGAGAAGCCAAAAATATTTTGGTAAACAAGATCAATGCTGGGATTTACTACTG 537
Db 478 TTTGTTGAAAAGCCCTGTGAGTTTGTATCCAAAGAAATTAATGCTGATTTACATCTCA 537
Qy 538 AACCCATCTGCTTGTGACCCGATTTGAGCTGAGGGCCACATCAATTGAGAAAGAGTCTTC 597
Db 538 AACCCCTTCCGTTCTTGTGACCTTATTAACCTCGTCTACCTCGATTTGAAAGGAAGTCTTT 597
Qy 598 CCTCAAAATTGCAGCTGATCAACAGCTCTATGCAATGGTTCACAGTTTTCAGGTTTTCGATGAT 657
Db 598 CCGCCATGTCATGACAGCAATTAACATCTTTCACACCTTGAGGGTTATTTGATGATGAT 657
Qy 658 GTTGTGACGCTAGGGACTACATTTACTTGGCTTGGCTCTTTATCTAGACTGATAGGAAG 717
Db 658 GTTGTGCAACCCCAAGGATTTATCTTACCCTGTTTGTGTTTATCTCTCTTGCCTGATG 717
Qy 718 AATCAGCTGCCAAGCTAG ---CTACTGGAGCAGCATGTTGTTGGCAATGCTGCTGCTGAT 774
Db 718 CATAAGCCTGAAATCTTTGGCCCCAGCTAGTAGCAATATCATCGGTAATGTTGATTTGAT 777
Qy 775 GAGAGCCCAAGATTGGAGAAAGTTGCTGATTTGGTCTCTGATGTCGCCATTTGGACCTGGG 834
Db 778 CCCTCAGCAACCATTTGCCAAAATTTGAAAATTTGTTCCCAATGTTGATTTGGTCCCAAC 837
Qy 835 TGTGTTTGTGAGGAGCGGCTGAGGCTTTCCCGCTGCACTGTGTCATCGCGGCGTGCATC 894
Db 838 GTAACGATTGCTGATGCTGTTGCTCTCAACGTTTGTGCCATTTCTTAAGTCTTCTCGCGTT 897
Qy 895 AGAAGCATGCTTGCATCTCAACAGCATTTATCGGCTGGGCACTCAACTGTTGGTCAATGG 954
Db 898 CGGACCATGCTCGGTTAAATCGAGTATTTAGGATGGAACCTACCCCTTTGGATCTTGG 957
Qy 955 GCAGGATAGAGATATGACTATCTCTGGGAGGATGTTTCATGTTGTGATGAGGTGTAC 1014
Db 958 AGTCGTTTGGAAAACGTTGCTGCTTAGGGGACGATGCTGTTGCAATGACGAGATTAC 1017
Qy 1015 AGCAATGGCGGTGTTGTTCTCCCAACAAGAGATCAAGTCAAGCATTTCTCAAGCCTG 1072
Db 1018 GTAAATGGTGGCAGTATTTTACCCTAAGAGCATTTAGTCTAACATTTGAGGTTCCCTG 1075
RESULT 12
AAH18598
ID AAH18598 standard; cDNA; 3220 BP.
XX AAH18598;
XX
XX 26-JUN-2001 (first entry)
XX Human cDNA sequence SEQ ID NO:18794.
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX Homo sapiens.
XX EP1074617-A2.
XX 07-FEB-2001.
XX 28-JUL-2000; 2000EP-0116126.
XX 29-JUL-1999; 99JP-0248036.
XX 27-AUG-1999; 99JP-0300253.
XX 11-JAN-2000; 2000JP-0118776.
XX 02-MAY-2000; 2000JP-0183767.
XX 09-JUN-2000; 2000JP-0241899.
XX (HELI-) HELIX RES INST.
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
DR

Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs -

Claim 8: SEQ ID 18794; 2537pp + CD ROM; English.

The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

Sequence 3220 BP: 631 A; 865 C; 1035 G; 689 T; 0 other:

[illegible]

Qy	541	CCATCTGTCCTTGCAGCGGCAATTCAGCTGACGCCAACATCATTTTGAGAAAGAGGTCTTCCCT	600
Db	763	CCTGCAGTGTCTGCGGCGCATCCAGCTGCAGCCTACGTCCATTGAGAAGAGAGGTCTTTCCCC	822
Qy	601	CAAATTGGCAGCTGATCAACAGACTCATATGCAATTCGTCTCTTTCCAGCTTTTTCATTCGGATGT	660
Db	823	ATTATGCCAACCAGGGGACGCTATATGCCATGGAGTTACAGGGCTTCTGGATGGACATT	882
Qy	661	GGTCAAGCCTTAGGACTACATTACTGGCTTTCGCTTTTATCTAGACTCGATTAGAAAGAAA	720
Db	893	GGGACAGCCAAAGGACTTCTCTACTGGCATGTGCCATGTCCCTTCTGCACTGAGAGCAGAAG	942
Qy	721	TCACTGCAAGCTACCTACCTACCTGACGACATATGTTTGGCAATGTGCTGGTCATGAGAGC	780
Db	943	CAGCCTTAGCGGCTGTGTCTCAGGCGCTTGGCCATTGTGGGAACGTGCTGGTGGACCAAG	1002
Qy	781	GCCAAAGATTGGAAGAGTGTCTGATTTGGTCTCATGTGCGCAATTCGCCACTGCTGCTGT	840
Db	1003	GCCCAGCATCGGCCCAAGCTGACATATGGCCCCAATGTGAGCTTGGGACTTGGCGTGGTG	1062
Qy	841	GTGGAGGACGGCTGAGGCTTTCGCGCTGCACTGTTCATCGCGGCGCTGGCTATCAACAAC	900
Db	1063	GTCGAAGATGCTGCTGTATCCGCGGCTCACAGGCTCTCGGGCATCCCGGATCCGTTCC	1122
Qy	901	CATGCTTGCATCTCAACAGCANATATCGGCTGGCACTCAACTGTTGGTCAATGGGACAG	959
Db	1123	CATTCTGGCTTGAAGTCTCTGCAATGTGGGCTGGCGCTGCCGCTGGGCTCAGTGGTAAG	1181
RESULT 13			
AAS41146			
ID	AAS41146	standard; cDNA; 990 BP.	
XX	XX		
AC	AAS41146;		
XX	XX		
DT	17-DEC-2001	(first entry)	
XX	XX		
DE	XX	cDNA encoding novel human enzyme polypeptide #362.	
XX	XX		
KW	Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase		
KW	ligase; hyperproliferative disorder; immunodeficiency disorder;		
KW	autoimmune disorder; neurological disorder; metabolic disorder;		
KW	inflammatory disorder; cardiovascular disorder; reproductive disorder;		
KW	blood-related disorder; infectious disorder; gene therapy; cytostatic;		
KW	anti arthritic; nephrotropic; anticoagulant; ss.		
OS	Homo sapiens.		
XX	XX		
PN	WO200155301-A2.		
XX	XX		
PD	02-AUG-2001,		
XX	XX		
Pf	17-JAN-2001; 2001MO-US01230.		
XX	XX		
PR	31-JAN-2000; 2000US-0179063.		
PR	04-FEB-2000; 2000US-0180628.		
PR	24-FEB-2000; 2000US-0184664.		
PR	02-MAR-2000; 2000US-0186350.		
PR	16-MAR-2000; 2000US-0189874.		
PR	17-MAR-2000; 2000VS-0190076.		
PR	18-APR-2000; 2000US-0198123.		
PR	19-MAY-2000; 2000US-0205515.		
PR	07-JUN-2000; 2000US-0209467.		
PR	28-JUN-2000; 2000US-0214880.		
PR	30-JUN-2000; 2000US-0215135.		
PR	07-JUL-2000; 2000US-0216647.		
PR	07-JUL-2000; 2000US-0216880.		
PR	11-JUL-2000; 2000US-0217487.		
PR	11-JUL-2000; 2000US-0217496.		
PR	14-JUL-2000; 2000US-0218290.		
PR	26-JUL-2000; 2000US-0220963.		
PR	26-JUL-2000; 2000US-0220964.		
PR	14-AUG-2000; 2000US-0224518.		

PR 14-AUG-2000; 2000US-02245119.
PR 14-AUG-2000; 2000US-02245213.
PR 14-AUG-2000; 2000US-02252114.
PR 14-AUG-2000; 2000US-02252114.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 18-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227182.
PR 30-AUG-2000; 2000US-0227099.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0232081.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234223.
PR 25-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0234998.
PR 27-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX
PI Rosen CA, Barash SC, Ruben SM;
XX WPI: 2001-465566/50.
DR P-PSDB; AAU23276.
XX
XX
PT Novel polypeptides and polynucleotides useful for diagnosing,
PT preventing, treating neural, immune system, muscular, reproductive,
PT pulmonary, cardiovascular, renal, proliferative disorders and cancerous
XX diseases -
PS Claim 4; SEQ ID No 372; 1180pp; English.
XX
XX The present invention relates to the isolation of novel human enzyme
CC polypeptides (AAU22915-AAU23814), and the cDNA and genomic sequences
CC encoding them. The enzyme polypeptides of the invention may comprise the
CC functional classes of oxidoreductases, transferases, hydrolases, lyases,
CC isomerases or ligases. The sequences of the invention are useful in the
CC diagnosis, treatment, prevention and/or prognosis of a wide range of
CC disorders including hyperproliferative disorders (e.g. cancer),
CC immunodeficiency disorders (e.g. AIDS) autoimmune disorders
CC (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease),
CC metabolic disorders (e.g. phenylketonuria), inflammatory disorders
CC (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis),
CC blood-related disorders (e.g. haemophilia), reproductive disorders
CC (e.g. infertility) and infectious disorders (e.g. Influenza). The
CC polynucleotides of the invention can also be used in gene therapy.

CC AAS40785-AAS41684 represent cDNA sequences encoding for the novel human
CC enzyme polypeptides of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 990 BP, 188 A, 272 C, 325 G, 196 T; 9 other:

Query Match 29.7%; Score 322; DB 22; Length 990;
Best Local Similarity 61.5%; Pred. No. 2.9e-92;
Matches 598; Conservative 3; Mismatches 342; Indels 5; Gaps 3;

QY 1 ATGACGCGCCCTGATTTGTCGGGGCTTTGCGGAACCCGCTTTCGCTTTTCACCTCTGAGC 60
DB 65 ATGACGCGCCTCTAGTTAGTTGGGGCTATGGGACGCGCTAGCGGCTGAGCGCTGAGC 124
QY 61 TTCGGAACCCCTCGTGGATTGTGCAACAGCCCATGATTCGACACGATGCGAGCT 120
DB 125 ACCCGCAAGCCATGCTGGCATTCGCAATGACCCATCTTCTGCACCAAGTTGGAGCG 184
QY 121 TTGAAGAAGATTGGGTCACAGAGGTGGTTTGGCTATCAACTATGCGCCGAGAGGTAATG 180
DB 185 CT-ANGCGCGAGCGCTGACACAGCTGATCCTGCGCTGAGCTACATCTGCGAGCTCTG 243
QY 181 ATTAATTTCTTGAAGACATTTGAGTAAGCTTTGGCATCACAAATTACATGCTCCCAAGAG 240
DB 244 GAGAGGAATGAAGCACAGGAGGAGGCTGGGAATCGGAATCCGAATCTCCATGTCCCATGAA 303
QY 241 ACTGAGCCCTTAGCAACCGCTGCGCTCTCTCTAGCAAGGACAAAGCTTGGCGGATGGA 300
DB 304 GAGGAGCCCTTTGGGACAGCTGGGCGCTTGGCGCTGCGCTGACTCTCTCTCTCTCTCT 360
QY 301 TCTGGCCAGCCATCTTGTCTCAACAGTGTGTCATGAAGCAATACCCATTTGCTGAA 360
DB 361 ACTGAGAGCCCTTCTCTGCTCTCAACAGTGTGTCATGAAGCAATACCCATTTGCTGAA 420
QY 361 CTCATCAAAATTCACAAAGTGTCTAGTGTGGTGAGGCAACAATATATGGTCACTAAGGTGGAT 420
DB 421 ATGTTGCAAGTTCACCGGCGCATGCGCAGGAGGGCTCCATCTGCTGACCAAGGTGGAG 480
QY 421 GAACCATCAAAATACCGTGTGTGGTGTGTCATGAAGCAAGCAAGCAAGTGTGCAAGAGTTT 480
DB 481 GAACCCCTCAATPACGGTGTGGTGTGTCATGAAGCAAGCAAGCAAGTGTGCAAGAGTTT 540
QY 481 GTTGAGAGCAAAATATTTGGTGTGTAAGATCAATGCTGGGATTTACTTACTTGAAC 540
DB 541 GTGGAAGCCACAGAGT 600
QY 541 CCATCTGCTCTTTCACCGCATTCAGCTGAGGCCAACATCAATTTGAGAAAGAGGTCTTCCCT 600
DB 601 CTGCAAGTGTGCGGCGCATCCAGCTGACGCTACGCTTCCATTCGAGAGAGGCTTCCCT 660
QY 601 CAATTCGAGCTGATCAACAGCTGTATGCAATGGTCTTCCAGGTTTGGATGGATGTT 660
DB 661 ATTATGCCAAGAGGGGCGCANNCTATGCTATGAGTGTACAGGGCTTCTGGATGGAGATT 720
QY 661 GGTGAGCTTAGGACTACATTTAGTGTGCTGCTTATCTATCTAGCTCGATTAGGAAGAAA 720
DB 721 GGGGAGCCCAAGAGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
QY 721 TCAGCTGCCAAGCTACTTACTTGAGCAGCATTTGTTTGGCAATGTCTTGGTGTGATGAGAG 780
DB 781 CACCTTGAGCGGCT-GTCTGAGCGCTGCGATTTGTGGGCAAGCTGCTGCTGCTGCTGCTGCT 839
QY 781 GCGAAGATTTGGAAGAGTTGCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
DB 840 GCCCGCATCGCCAGAACTGACGATTTGCCCAATGTGAGCCCTGGACCTGCGCGGTG 899
QY 841 GTGGAGAGCGGCTGAGGCTTTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
DB 900 GTGGAAGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 959
QY 901 CATGCTTG 908

DB 960 CATTCCTG 967
RESULT 14
AAF13173
ID AAF13173 standard; cDNA; 1351 BP.
XX
AC AAF13173;
XX
DT 13-MAR-2001 (first entry)
XX
DE Aspergillus oryzae EST SEQ ID NO:5696.
XX
KW Multiple gene expression; filamentous fungal cell; EST;
KW expressed sequence tag; Fusarium venenatum; Aspergillus niger;
KW Aspergillus oryzae; Trichoderma reesei; identification; recombination;
KW culture condition; environmental stress; spore morphogenesis;
KW metabolic pathway engineering; catabolic pathway engineering; ss.
XX
OS Aspergillus oryzae.
XX
PN WO200056762-AZ.
XX
PD 28-SEP-2000.
XX
PF 22-MAR-2000; 2000WO-US07781.
XX
PR 22-MAR-1999; 99US-0273623.
XX
PA (NOVO) NOVO NORDISK BIOTECH INC.
XX (NOVO) NOVO NORDISK AS,
XX Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;
XX WPI; 2000-594572/56.
DR
XX
PT Monitoring differential expression of genes in filamentous fungal cells
PT uses fluorescence-labeled nucleic acids isolated from the cells and a
PT substrate of expressed sequence tags -
XX
PS Claim 88; Page 2361; 3161pp; English.
XX
CC The present invention describes a method for monitoring differential
CC expression of genes in a first filamentous fungal (FF) cell relative to
CC expression of the same genes in one or more second filamentous fungal
CC cells. The method uses fluorescence-labeled nucleic acids isolated from
CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs
CC are used in the methods for monitoring differential expression of genes
CC in a first filamentous fungal (FF) cell relative to expression of the
CC same genes in one or more second filamentous fungal cells. Monitoring
CC the global expression of genes from FF cells allows the production
CC potential of the microorganisms to be improved. New genes may be
CC discovered, possible functions of unknown open reading frames can be
CC identified and gene copy number variation and stability can be
CC monitored. The expression of genes can be used to study how FF cells
CC adapt to changes in culture conditions, environmental stress, spore
CC morphogenesis, recombination, metabolic or catabolic pathway
CC engineering. Using ESTs provides several advantages over genomic or
CC random cDNA clones including elimination of redundancy as one spot on an
CC array equals one gene or open reading frame, and Organisation of the
CC microarrays based on function of the gene products to facilitate
CC analysis of the results. AAF07478 to AAF11247 represents ESTs from
CC Fusarium venenatum; AAF11248 to AAF11853 represents ESTs from Aspergillus
CC niger; AAF11854 to AAF14878 represents ESTs from Aspergillus oryzae; and
CC AAF14879 to AAF15337 represents ESTs from Trichoderma reesei, which are
CC all specifically claimed in the present invention.
XX
SQ Sequence 1351 BP; 305 A; 365 C; 315 G; 365 T; 1 other;

Query Match 27.9%; Score 300.4; DB 21; Length 1351;
Best Local Similarity 57.5%; Pred. No. 2.9e-85;
Matches 604; Conservative 0; Mismatches 431; Indels 15; Gaps 3;

QY 49 TTGACTCTGAGCTTCCCGAAGCCCTCGTGGATTTTGCAGAAAGCCCAAGATTCGTGCAC 108
Db 9 TAGACATTGACCTTCCCAAGCCCTCGTGGAAATTTGGGAACCGTCCCATGATTTTGCAC 68
QY 109 CAGATCGAAGCTTTGAAGAAAGTTGGGTGCACAGAGGTGGTTTGGCTATCAACTATCGC 168
Db 69 CAAGTCGAGAGCTTGGCTGCTGGTGTACGGGATTTTGTCTGGCTGTTAACTACCGC 128
QY 169 CCAGAGTAATGATTAATTTCTTTGAAGGACTTTTGAAGATAAGCTTGGCATCACAATTACA 228
Db 129 CCAGATGTCATGTTTCGGCCCTCAAGAGATATGAGAACATACATGTAATGTGAATCGAG 188
QY 229 TGCTCCCAAGAGACTGAGCCCTTAGGAACCGCTGCGCCCTCTTCTCTAGCAAGGGACAAG 288
Db 189 TTTCTCCGTCCGAGTCGGAACCCCTGGGCACTGCGCGTCCCTGAAAGTTAGC---GGAAAG 245
QY 289 CTTCGGGATGATTCGGCCAGGCATCTTTGTCTCAACAGTGTATGATCAAGCGAATAC 348
Db 246 ATTTTGGGCAAGGACGACTCTCCATTTCTGTTCTTAAACTCCGATGTCATCTCGGACTAC 305
QY 349 CCATTGTCTGAACATCAATAATTTCCAAAGTGTCTGTTGGTGGAGCAACAATTATGGTC 408
Db 306 CCTTCAAGGAGCTGGCTGAGTTCCATTAAGAGCATGGCGATGAAGGTACCATTTGTTGT 365
QY 409 ACTAAGTGGATGAACCATCAAAATACGGTGTGTGGTT---ATGGAGGAGGCAACTGGC 465
Db 366 ACTAAGTTCGAGGAGCCCTCCAAAGTATGCGGTGCTGTTGTTCACAAGCCCAATCAACCCCTCG 425
QY 466 AGGTGGAAAGTTTGTGTGAGAGCCCAAAATATTTGTGGGTGAACAAGATCAATGCTGGG 525
Db 426 CGCATCTGACCGCTTGTGCGAAAGCCTGTTGAATTCGTGTGAAACCGCATCAACGCCGT 485
QY 526 ATTTACTACTGAACCCATCTGCTGACCCATTTGACCCATGAGCTGAGGCCCAACATCAATTGAG 585
Db 486 ATCTACATCATGAACCCCTAGTGTGTGAATCGCATGAGCTGCGGCCGACATCCCATTTGAG 545
QY 586 AAAGAGTCTTCCCTCAAAATGCACTGATCAACAGCTATGCAATGGTCTCTCCAGGT 645
Db 546 CAGGAGACATTCGCGGCTATCTGCAAGGACGACAGCTCCATCTCTTGAACCTGGAGGT 605
QY 646 TTTGGATGGATTTGGTTCAGCTAGGACTACATTAATGCTGGTGTGGTCTTTATCTAGAC 705
Db 606 TTTGGATGGAGTTGGTTCAGCTAGGACTTCTCAGCGGCACCTGTCTATACCTCACT 665
QY 706 TCGATTAGGAAGAAATCAGCTCCCAAGCTAGCTACTGAGGACACATGTT-----GTT 756
Db 666 TCTCTCGCAAGCGCAATTCGAAGCTGCTCGCACCCCACTCCGAACCCCTACGTGTACGGT 725
QY 757 GGCAATGTGCTGGTGCATGAGAGCCCAAGATTGGAGAAGTTGTCTGATTTGGTCTGTAT 816
Db 726 GGCAATGTATGGTTGACCCCTCCGCCAAGATTGGCAAGAATTTGGCGCATTTGGTCCCAAT 785
QY 817 GTCGCCATTTGACCTGGGTGTGTTGTGGAGGACGCGTGTAGGCTTTTCCGCTGCACTGTC 876
Db 786 GTAGTCATTTGTCACCAAGTTGTGTGCTGATGAGTTCTGCTGAGGCTGTGTCAGGTTGTGCTT 845
QY 877 ATGCGCGGCTGCTGATCAAGAGCATGCTTGCATCTCAAGAGCATATATCGGTGGGCAC 936
Db 846 CTAGAGAACAGCAAGGTCAAGACCATGCTGGGTTAAGTCAACCATTTGTCGGATGGAAT 905
QY 937 TCACTGTTGTTCATGGCGGATAGAGATATGACTATCTCTGGGGAGGATGTTTAT 996
Db 906 AGCTCGTAGGAGGTGGGCTGCTTTTGGAGAATGTACCGTGTGGGTGACGACGTACT 965
QY 997 GTGTGTGATGAGGTACAGCAATGGCGGTGTTGTCTCCCAACATAAAGAGATCAAGTCA 1056
Db 966 ATCGACAGCAAGTCTATGTTAATGGTGGATCCATCCCTCTCATAGAGCATCAAGCAG 1025
QY 1057 AGCATTCGAAGCCCTGAGATCGTCAATGTA 1086
Db 1026 AACGTTGATGTTCCCGCCATTTATTATGTA 1055

RESULT 15
AAC43567
ID AAC43567 standard; DNA; 458 BP.
XX
AC AAC43567;
XX
DT 17-OCT-2000 (first entry)
XX
DE Zea mays DNA fragment SEQ ID NO: 39699.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic;
KW pathway; promoter; termination sequence; corn; ss.
XX
OS Zea mays subsp. mays.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.

```
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146380.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147202.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.

PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 24.2%; Score 262.6; DB 21; Length 458;
Best Local Similarity 96.8%; Pred. No. 2e-73;
Matches 268; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 ATGAAGCCCTCATTTCTCGGGGGTTTCGAAACCCGCTTCGCGCTTTTCACTCTGACC 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 182 ATGAAGCCGCTCATTTCTCGGGGGTTTCGAAACCCGCTTCGCGCTTTTCACTCTTAGC 241
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 TTCCCGAAACCCCTCGTGGATTTTGCAAAAGCCCATGATTCGCACCATCAAGCT 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 242 TTCCCGAAAGCCCTTCTCGATTTTCCAAAACAGCCCATGATTCGCATCACTCAAGCT 301
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 TTCAAAAGACTTTGGGGTTCACAGAGGTGGTTTTTGGCTATCAACTATCGCCAGAGTAATG 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 302 TTGAAGAGATTGGGGTTCACAGAGGTGGTTTTTGGCCATCACTACCGCCAGAGTAATG 361
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: November 27, 2002, 02:46:27
Job time : 264 secs


```
Query Match          6.8%; Score 74.2; DB 3; Length 1535;
Best Local Similarity 49.5%; Pied. NO. 4e-14;
Matches 255; Conservative 0; Mismatches 248; Indels 12; Gaps 2;

QY 31 GGAACCGCGCTTCGGCTTGTGACTCTGAGCTCCCGAAGCCCTCGTGGATTGTCGAAC 90
DB 113 GGAACCGCGCTTCAGACTTGTCTTTGAGGTGCCCAACCATTTGCTCGTGGCAGGG 172

QY 91 AAGCCCATGATCTGCACACGATCGAAGCTTTGAAGAAGTTGGGGTCAAGAGAGTGGTT 150
DB 173 GTCCCTATGATCAACACCATATTAAGCGCTTGCCAGGTCCTGGAATGCGAGAGATT 232

QY 151 TTGGCTATCAACATATCGCCCA-----GAGGTAATGATTAATTTCTTGAAGACTTTGAC 204
DB 233 CTGCTCATTTGGCTTCTTACCAACCTGATGAGCCCTCACCCAGTTCTTAGAAGCGCCAG 292

QY 205 GATAAGCTTGGCATCAATATGATGCTCCCAAGAGACTGAGCCCTTAGAAGCGCTGGC 264
DB 293 CAGGAGCTTTAAACCTTCCACTCAGGTACCTGCAGGAATTTGCCCTTAGCAGAGGGGT 352

QY 265 CTTCTTGGCTCTAGCAGGAGCAAGCTTGGCGATGGATCTGGCCAGCCATTCTTTGTCCTC 324
DB 353 GGTCTTTAGCATTTTCGAGACAGATCTCGGTGGGAGCCCGAGGCATTTCTCGCTC 412

QY 325 AACACGATGCTATACAGGCAATACCAATTTGCTGAAGTCAATCAAAATTTCAAAAGTCTCAT 384
DB 413 AATGCTGATGTCTGCTCCGACTTCCCTTTGAGTGCTATGTTGGAAGCCCAAGCGCCAG 472

QY 385 GGTGGTGAGCAACATATGTCACATAAG-----TGGATCAACGATCAAAATACGGT 438
DB 473 CTTCAACCTTCTTACTCTCTGCACTAGCGCTTAACAGCAGCAATCCCTCAACTACGGC 532

QY 439 GTTGTGTTATGAGGAGGCACTGCGAGGTGGATGTTTCTTGAGAGCCCAAAATA 498
DB 533 TGCATCGTTGAGAAATCCAGACACAGAGGATATGCACTATGTGGAGAAACCCAGACA 592

QY 499 TTTGTGGTAAACATCAATGTTGGGATTTACTT 533
DB 593 TTTATCAGTGACATCACTACCTGCGGCATCTACCT 627

RESULT 2
US-09-007-119-20/c
; Sequence 20, Application US/09007119C
; Patent No. 6300541
; GENERAL INFORMATION:
; APPLICANT: Lighfoot, David A.
; APPLICANT: Gibson, Paul T.
; APPLICANT: Merkem, Khalid
; TITLE OF INVENTION: Soybean Sudden Death Syndrome Resistant Soybeans,
; TITLE OF INVENTION: Soybean Cyst Nematode Resistant Soybeans and Methods of
; TITLE OF INVENTION: Breeding and Identifying Resistant Plants
; FILE REFERENCE: Sou Illinois 1268/2 Sequence Listing
; Patent No. 6300541
; CURRENT APPLICATION NUMBER: US/09/007,119C
; CURRENT FILING DATE: 1998-01-14
; EARLIER APPLICATION NUMBER: 60/035,335
; EARLIER FILING DATE: 1997-01-14
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 20
; LENGTH: 787
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3)
; OTHER INFORMATION: (a or c or g or t/u)
```

```
; NAME/KEY: misc_feature
; LOCATION: (400)..(401)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (421)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (428)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (432)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (462)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (475)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (514)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (521)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (526)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (532)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (542)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (572)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (574)..(575)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (589)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (603)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (609)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (614)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (636)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
```

```
; LOCATION: (638)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (671)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (687)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (696)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (703)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (708)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (712)
```

Query Match 6.7%; Score 73.2; DB 4; Length 787;

Best Local Similarity 56.0%; Pred. No. 5.7e-14;

Matches 227; Conservative 0; Mismatches 151; Indels 27; Gaps 5;

```
QY 524 GGATTTACTTGAACCCATCTCTGACCGCATTTGAGCTGAGCGCAACATCAATTG 583
    ||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 563 GGGTAACTGGATGAACCCCTTNGGTATGGANAGATNGAGNGACACCATTTCTTGA 504
QY 584 AGAAGAGGTCTCCCTCAAAATGTCAGCTGATCAACAGCTCTATGCAATGGTCTTCC-- 641
    ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 503 AGACAGAGGGTTTCCAAATATTTTCNGATAGAAAGCTGNATGCAATGGTCTTCCAC 444
QY 642 ----AGGTTTTGGATGATGTGGTCAGCCTAGGAGCTACATTAATGCTGCGTCTTT 697
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 443 AGGAGGATTTNAGANGAAATNGGACAA---AGGGACTATATTTNNGANTGAGACTCN 387
QY 698 ATCTAGACTCGATTAGGAAGA---AATCAGCTGCCAGTAGTACTTGGAGCATGTGTG 754
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 386 ACNTGAACCTGGGAGGAATAATTTCTTAGTGAACNGCCCAATGACNTCANACATNTG 327
QY 755 TTGGCAATCTGTGTGTCATGAGAGCGCAAGATTGGAGAAGTTGTCTGATTTGGTCCG 814
    ||| ||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 326 TGGGAAATCTGTTGTGTCATGAGACTGCCAA-ATNGGTGATGTCATGCAAAATGGTCC-- 270
QY 815 ATGTCGCCATTGGACCTGGTGTGTGTGGAGGACGCGTGAGGCTTCCCGCTGCACTG 874
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 269 -----ATCCNGGATGTATCATTTGAGAAAGAGTTGGGCTCAAAATCTNACACTG 222
QY 875 TCATGCGCGCGTGCATATCAAGAAGCATCTTGCATCTCAAAACA 919
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 221 TTATGTGAAGCACTCGTGTATAAAGCAATTTGTTATATCNGACA 177
```

RESULT 3

US-08-961-527-112

; Sequence 112, Application US/08961527

; Patent No. 6420135

; GENERAL INFORMATION:

; APPLICANT: Charles Kunsch

; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences

; NUMBER OF SEQUENCES: 391

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville

; STATE: Maryland

; COUNTRY: USA

;/ ZIP: 20850
;/ COMPUTER READABLE FORM:
;/ MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
;/ COMPUTER: HP Vectra 486/33
;/ OPERATING SYSTEM: MSDOS version 6.2
;/ SOFTWARE: ASCII TEXT
;/ CURRENT APPLICATION DATA:
;/ APPLICATION NUMBER: US/08/961,527
;/ FILING DATE:
;/ CLASSIFICATION: 424
;/ PRIOR APPLICATION DATA:
;/ APPLICATION NUMBER:
;/ FILING DATE:
;/ ATTORNEY/AGENT INFORMATION:
;/ NAME: BROOKES, A. Anders
;/ REGISTRATION NUMBER: 36,373
;/ REFERENCE/DOCKET NUMBER: PB340P1
;/ TELECOMMUNICATION INFORMATION:
;/ TELEPHONE: (301) 309-8504
;/ TELEFAX: (301) 309-8512
;/ INFORMATION FOR SEQ ID NO: 112:
;/ SEQUENCE CHARACTERISTICS:
;/ LENGTH: 7902 base pairs
;/ TYPE: nucleic acid
;/ STRANDEDNESS: double
;/ TOPOLOGY: linear
;/ US-08-961-527-112

Query Match 5.4%; Score 58.4; DB 4; Length 7902;
Best Local Similarity 52.5%; Pred. No. 1.6e-08;
Matches 128; Conservative 0; Mismatches 116; Indels 0; Gaps 0;
QY 2 TGAAGGCCCTCATCTTGTGCGGGGTTTCGGAACCGGCTTCGGCCTTTGACTCTGAGCT 61
Db 4961 TGAAGGCCATTTATCTAGCAGCGGGATTTCGGAACCTGCTTCGCTATGACTGAAAATA 5020
QY 62 TCCCGAAACCCCTCGTGGATTGTCGAACAGCCCATGATTCGACACAGATCGAAGCTT 121
Db 5021 CCCCTAAAGCCTTGGTTGAGTTAATCAAAAAGCCTTTGATTGAGTACGAAATTTGACTTTC 5080
QY 122 TCAAGAACTTGGGTCACAGAGCTGCTTTGGCTATCAACTATCCGCCACAGAGTAAATGA 181
Db 5081 TCAAGAAAGGAATCAATGACATCATCATTTGTTGTTATCTTTAAGAGACAATTCG 5140
QY 182 TTAATTTCTTGAAGGACTTTGAGGATAAGCTTGGCATCAATACATGCTCCCAAGAGA 241
Db 5141 ATTACTTGAAGAGAAATFACGGTGTTCGTCCTGTTTCAATGATAAATACGCTGACTACA 5200
QY 242 CTGA 245
Db 5201 ATAA 5204

RESULT 4
US-09-024-023-1
; Sequence 1, Application US/09024023
; Patent No. 6110899
; GENERAL INFORMATION:
; APPLICANT: Lonetto, Michael A.
; TITLE OF INVENTION: liec
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert, Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2793
; COMPUTER READABLE FORM:
; MEDIUM TYPE: diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0

;/ CURRENT APPLICATION DATA:
;/ APPLICATION NUMBER: US/09/024,023
;/ FILING DATE:
;/ CLASSIFICATION:
;/ PRIOR APPLICATION DATA:
;/ APPLICATION NUMBER: 60/039,210
;/ FILING DATE: 28-FEB-1997
;/ ATTORNEY/AGENT INFORMATION:
;/ NAME: Falk, Stephen T
;/ REGISTRATION NUMBER: 36,795
;/ REFERENCE/DOCKET NUMBER: GM50018
;/ TELECOMMUNICATION INFORMATION:
;/ TELEPHONE: 215-994-2488
;/ TELEFAX: 215-994-2222
;/ TELEX:
;/ INFORMATION FOR SEQ ID NO: 1:
;/ SEQUENCE CHARACTERISTICS:
;/ LENGTH: 690 base pairs
;/ TYPE: nucleic acid
;/ STRANDEDNESS: double
;/ TOPOLOGY: linear
;/ US-09-024-023-1

Query Match 5.1%; Score 55.2; DB 3; Length 690;
Best Local Similarity 51.6%; Pred. No. 4.3e-08;
Matches 126; Conservative 0; Mismatches 116; Indels 0; Gaps 0;
QY 2 TGAAGGCCCTCATCTTGTGCGGGGTTTCGGAACCGGCTTCGGCCTTTGACTCTGAGCT 61
Db 2 TGAAGGCCATTTATCTAGCAGCGGGATTTCGGAACCTGCTTCGCTATGACTGAAAATA 61
QY 62 TCCCGAAACCCCTCGTGGATTGTCGAACAGCCCATGATTCGACACAGATCGAAGCTT 121
Db 62 CCCCTAAAGCCTTGGTTGAGTTAATCAAAAAGCCTTTGATTGAATACCAAAATTCAGTTTC 121
QY 122 TGAAGAACTTGGGTCACAGAGTGGTTTGGCTATCAACTATTCGCCACAGAGTAAATGA 181
Db 122 TCAAGAAAGGAATCAATGACATCATCATCTGTTGGTTATCTTTAAAGAGACAATTCG 181
QY 182 TTAATTTCTTGAAGGACTTTGAGGATAAGCTTGGCATCAATTAATGCTCCCAAGAGA 241
Db 182 ATTATTTAAAGAAATAATACGGTGTTCGCTCGTTTCAATGATAAATACGCTGACTACA 241
QY 242 CTGA 245
Db 242 ATAA 245

RESULT 5
US-09-531-111-1
; Sequence 1, Application US/09531111
; Patent No. 6303571
; GENERAL INFORMATION:
; APPLICANT: Lonetto, Michael A.
; TITLE OF INVENTION: liec
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert, Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2793
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/531,111
; FILING DATE: 17-Mar-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:

```

; REGISTRATION NUMBER: 36,795
; REFERENCE/DOCKET NUMBER: GM50018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2488
; TELEFAX: 215-994-2222
;
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 287 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
US-09-024-023--3

Query Match          4.9%; Score 53.6; DB 3; Length 287;
Best Local Similarity 51.2%; Pred. No. 8.5e-08;
Matches 125; Conservative 0; Mismatches 119; Indels 0; Gaps

QY  2 TGAAGGCCCTCATCTTGTTCGGGGGTTTCGGAACCCGCCCTTCGGCCTTTGACTCTGAGCT 61
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  62 TCCGGAACCCCTCGTGATTTGCAACAAGCCCATGATTCGCACCAGATCGAAGCTT 121
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  62 CCCCTAAAGCCTTGGTTCAGGTTAATCAAAAGCCTTTGATTGAATACCAAATTCGATTC 121
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  122 TGAAGAAGTGTGGGTCACAGAGTGCTTTTGGCTATCAACTATGCCACAGAGGTAATGA 181
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  182 TTAATTTCTTGAAGGACTTTGAGGATAAGCTTGGCATCACAATTACATGCTCCCAAGAGA 241
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  242 CTGA 245
Db   || |

QY  242 ATAA 245
Db   || |

RESULT 7
US-09-531-111-3
; Sequence 3, Application US/09531111
; Patent No. 6303571
; GENERAL INFORMATION:
; APPLICANT: Lonetto, Michael A.
; TITLE OF INVENTION: licC
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert, Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2793
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq For Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/531,111
; FILING DATE: 17-Mar-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/024,023
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Falk, Stephen T
; REGISTRATION NUMBER: 36,795
; REFERENCE/DOCKET NUMBER: GM50018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2488
; TELEFAX: 215-994-2222

```

```
/
/
/ TELEX: <Unknown>
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 287 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-531-111-3

Query Match          4.9%; Score 53.6; DB 4; Length 287;
Best Local Similarity 51.2%; Pred. No. 8.5e-08;
Matches 125; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

QY  2  TGAAGCCCTCATCTTGTGCGGGGTTTCGCAACGCCCTTCGCGCTTTCACCTGAGCT 61
DB  2  TGAAGCCATATCTTAGCAGCGGGATTGGAACTCGCTTGCCTATACACTGAAATA 61

QY  62  TCCGAAACCCCTCGTGGATTTCGAAACAGCCCATGATCTCGCACAGATCGAAGCTT 121
DB  62  CCCTAAAGCCTTGGTTACGTTTAATCAAAACCTTTGATGAATACCAAAATTCGATTC 121

QY  122 TGAAGAAGTTGGGTCACAGAGGTGGTTTGGCTATCAACTATCGCCACAGAGGTAATGA 181
DB  122 TCAAGAAAAAGGAATCAATGACATCATCATCGTTGGTTATCTTAAAGAAACAATTCG 181

QY  182 TTAATTTCTTGAGGACTTTCAGCATAAAGCTTGGCATCAAAATACATGTCCTCCACAGA 241
DB  182 ATTAATTAAGAAAAAATACGGTGTTCGGCTCGTTTTCATGATGAATAATACGCTGACTACA 241

QY  242 CTGA 245
DB  242 ATAA 245

RESULT 8
US-08-196-218-31
; Sequence 31, Application US/08196218
; Patent No. 5614619
; GENERAL INFORMATION:
; APPLICANT: Piepersberg, Wolfgang
; APPLICANT: Stockmann, Michael
; APPLICANT: Taleghani, Kampiz Mansouri
; APPLICANT: Distler, Jurgen
; APPLICANT: Grabley, Susanne
; APPLICANT: Sichel, Petra
; APPLICANT: Brau, Barbara
; TITLE OF INVENTION: Secondary-Metabolite Biosynthesis Genes
; TITLE OF INVENTION: From Actinomycetes, Method of Isolating Them, and Their
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSER: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/196,218
; FILING DATE: 25-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ogden, Stasia L.
; REGISTRATION NUMBER: 36,228
; REFERENCE/DOCKET NUMBER: 02481.1372-00000
; TELECOMMUNICATION INFORMATION:
```

```
/
/
/ TELEPHONE: 202-408-4000
/ TELEFAX: 202-408-4400
/ INFORMATION FOR SEQ ID NO: 31:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2634 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 3..401
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 416..1531
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 1561..2625
US-08-196-218-31

Query Match          4.9%; Score 52.8; DB 1; Length 2634;
Best Local Similarity 57.1%; Pred. No. 5.8e-07;
Matches 96; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY  1  ATGAAGCCCTCATCTTGTGCGGGGTTTCGGAACCCGCCCTTCGCCCTTTCAGCTCTGAGC 60
DB  1561 ATGAAGCTCTGGTGCTCGCGCGGGATCTGTTACCCGCTTGGGCTTTCAGTTATTTCG 1620

QY  61  TTCCGAAACCCCTCGTGGATTTCGCAACAGCCCATGATTCGACACAGATCGAAGCT 120
DB  1621 ATGCCAAACAACTGATCCCATGCCCAACACACCCGCTGCTGCTCATGCTGCAACGCC 1680

QY  121 TTGAAGAAGTTGGGTCACAGAGGTGGTTTGGCTATCAACTATCGC 168
DB  1681 GTCCGGAGCTGGCGGTGACCGAGGTGCGGTGCTCATGCTCGCAACCCG 1728

RESULT 9
US-08-681-953-31
; Sequence 31, Application US/08681953
; Patent No. 5710032
; GENERAL INFORMATION:
; APPLICANT: Piepersberg, Wolfgang
; APPLICANT: Stockmann, Michael
; APPLICANT: Taleghani, Kampiz Mansouri
; APPLICANT: Distler, Jurgen
; APPLICANT: Grabley, Susanne
; APPLICANT: Sichel, Petra
; APPLICANT: Brau, Barbara
; TITLE OF INVENTION: Secondary-Metabolite Biosynthesis Genes
; TITLE OF INVENTION: From Actinomycetes, Method of Isolating Them, and Their
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSER: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/681,953
; FILING DATE: 30-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/196,218
; FILING DATE: 25-AUG-1994
```

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Ogden, Stasia L.
; REGISTRATION NUMBER: 36,228
; REFERENCE/DOCKET NUMBER: 02481.1372-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2634 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..401
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 416..1531
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1561..2625
; US-08-681-953-31

Query Match 4.9%; Score 52.8; DB 1; Length 2634;
Best Local Similarity 57.1%; Pred. No. 5.8e-07;
Matches 96; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

Qy 1 ATGAAGGCGCTCATCTTGTGCGGGGGTTTCGGAACCGCGCTTCGGCCTTTGACTCTGAGC 60
Db 1561 ATGAAGGCTGTGGTCTGCGCGCGGATCTGTGATCCCGCTGCGGCTTTCAGTTATTGG 1620

Qy 61 TTCGGAAACCCCTCGTGGATTTTCAACACAGCCCATGATCTGCACACGATCGAAGCT 120
Db 1621 ATGCCCAACCACTGATCCCATCGCAACACACACCGCTGCTGGTGATGCTGAACGCC 1680

Qy 121 TTGAAGAAGTTGGGTCACAGAGTGGTTTGGCTATCAACTATCGC 168
Db 1681 GTCCGGAGCTGGGCTGACCGAGGTCGGCTCATCGTCGGCAACGCC 1728

RESULT 10
US-09-370-838-155
; Sequence 155, Application US/09370838
; Patent No. 6444425
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Roadoh
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
; TITLE OF INVENTION: LUNG CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.475C1
; CURRENT APPLICATION NUMBER: US/09/370,838
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/285,323
; EARLIER FILING DATE: 1999-04-02
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 155
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Homo sapien
; US-09-370-838-155

Query Match 4.3%; Score 46.2; DB 4; Length 678;
Best Local Similarity 49.5%; Pred. No. 3.8e-05;
Matches 151; Conservative 0; Mismatches 148; Indels 6; Gaps 1;

Qy 235 CAAGAGACTGAGCCCTTAGGAACCGCTTGTCTAGCAAGGACCAAGCTTGGC 294
Db 350 CAGGAATTTGCCCCCTTAGGCACAGGGGGTGTCTTTACCATTTTCGAGACCAGATCCTG 409
```

```

Qy 295 GATGGATCTGGCCAGCCATTTCTTGTCTCTCAACAGTGAATGTCATAAGGCAATACCCATT 354
Db 410 GCTGGGAGCCCGAGGCAATTTCTGCTCAATGCTGATGCTGCTCGACTTCCCTTG 469
Qy 355 GCTGAACATCATCAAAATTTCAACAAGTGTCAATGGTGTGAGGCAACAATATGTCACCTAAG 414
Db 470 AGTGCTATGTTGGAAGCCCGGACGCGTCAACCTTTTCTTACTCCTTTGGCACTACG 529
Qy 415 GTGGATG-----AACCATCAAAATACGGTGTGTTGGTTATGGAGGCAACTGGCAGG 468
Db 530 GCTAACAGGAGGCATATCCCTCACTACGCGTGCATCGTTGAGATCCACACACACGAG 589
Qy 469 GTGGAAGGTTTGTGGAAGCCAAAAATATTTTGGTGTAAACAAGATCAATGCTGGGATT 528
Db 590 GTATTGCACATATGTGGAGAAACCCAGCACATTTATCATGTGACATCATCAACTGCGGCACC 649
Qy 529 TACTT 533
Db 650 TACCT 654

RESULT 11
US-09-194-905-7
; Sequence 7, Application US/09194905
; Patent No. 6306627
; GENERAL INFORMATION:
; APPLICANT: DECKER, Heinrich
; TITLE OF INVENTION: ISOLATION OF THE BIOSYNTHESIS GENES FOR
; TITLE OF INVENTION: PSEUDO-OLIGOSACCHARIDES FROM STREPTOMYCES GLAUCESCENS
; TITLE OF INVENTION: GLA.O AND THEIR USE
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY & LARDNER
; STREET: 3000 K Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/09/194,905
; APPLICATION NUMBER: US/09/194,905
; FILING DATE: 29-JUL-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/EP97/02826
; FILING DATE: 30-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 19622783.6
; FILING DATE: 07-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Granados, Patricia D.
; REGISTRATION NUMBER: 33,683
; REFERENCE/DOCKET NUMBER: 026083/0193
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6854 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-09-194-905-7

Query Match 3.8%; Score 41.6; DB 4; Length 6854;
Best Local Similarity 55.6%; Pred. No. 0.0049;
Matches 80; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
```



```
; FILING DATE: 435
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/377,440
; FILING DATE: 24-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: GOLDBERG, JULES E.
; REGISTRATION NUMBER: 24,408
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-986-4090
; TELEFAX: 212-818-9479
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28804 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FRAGMENT TYPE: N-terminal
; US-08-592-874-1

Query Match          3.7%; Score 40.2; DB 2; Length 28804;
Best Local Similarity 45.0%; Pred. No. 0.033;
Matches 150; Conservative 0; Mismatches 183; Indels 0; Gaps 0;

QY 1 ATGAAGCCCTCATCTTGTGCGGGGTTTCGGAACCCGCTTCGGCCCTTTGACTCTGAGC 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23238 ATGAAGGTATCATCTTTCGGGGGCGAGCGGCGCTCTACCCGCAACGCTGTCG 23297

QY 61 TTCCGAAACCCCTCGTGGATTTCGAAACAGCCCATGATCTGCACCAAGATCGAAGCT 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23298 ATCTCGAACAGCTGCTTCCCGTCTATGACAAGCCCATGATCTTACCCCTGTCGGTG 23357

QY 121 TTGAAAGAAAGTTGGGTCACAGAGTGGTTTGGCTATCAACTATGCCCGCAGAGTAATG 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23358 CTGATGCTCAGGGTATCCGGGACATCTTGATCATCTCCACCCCGCGACCTGCCGATG 23417

QY 181 ATTAATTTCTGAAGACTTTGAGGATAAGCTTGGGATACACATTTACATGCTCCCAAGAG 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23418 TTCCAGGCGCTGCTCGGCGACGTTTCGGCATTCGGCATCAACCTGAGCTATGCCGAACAG 23477

QY 241 ACTGAGCCCTTAGAACCCGCTGCCCTCTTGCTCTAGCAAGGACAAAGCTTGGCGGATGA 300
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23478 CCTTCGCCCAACGGCTTCGGAAGCCTTCATCATCGGCGCGGATTTTCGTCGCAACGAT 23537

QY 301 TCTGGCCAGCATCTTTTGTCTCAACAGTGAT 333
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23538 CCCAGCGCTGATCTTCGGGCAACATCTAT 23570

RESULT 15
US-09-096-867-2
; Sequence 2, Application US/09096867
; Patent No. 6030817
; GENERAL INFORMATION:
; APPLICANT: Pollock, Thomas J
; APPLICANT: Mikolajczak, Marcia
; APPLICANT: Yamazaki, Motohide
; APPLICANT: Thorne, Linda
; APPLICANT: Armentrout, Richard W
; TITLE OF INVENTION: Production of Xanthan Gum by Sphingomonas Bacteria
; FILE REFERENCE: seq list for app'l filed from pro. appl
; CURRENT APPLICATION NUMBER: US/09/096,867
; CURRENT FILING DATE: 1998-06-11
; EARLIER APPLICATION NUMBER: 60/049,428
; EARLIER FILING DATE: 1997-06-12
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 28804
; TYPE: DNA
; ORGANISM: Sphingomonas sp. S88
US-09-096-867-2

Query Match          3.7%; Score 40.2; DB 3; Length 28804;
Best Local Similarity 45.0%; Pred. No. 0.033;
Matches 150; Conservative 0; Mismatches 183; Indels 0; Gaps 0;

QY 1 ATGAAGCCCTCATCTTGTGCGGGGTTTCGGAACCCGCTTCGGCCCTTTGACTCTGAGC 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23238 ATGAAGGTATCATCTTTCGGGGGCGAGCGGCGCTCTACCCGCAACGCTGTCG 23297

QY 61 TTCCGAAACCCCTCGTGGATTTCGAAACAGCCCATGATCTGCACCAAGATCGAAGCT 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23298 ATCTCGAACAGCTGCTTCCCGTCTATGACAAGCCCATGATCTTACCCCTGTCGGTG 23357

QY 121 TTGAAAGAAAGTTGGGTCACAGAGTGGTTTGGCTATCAACTATGCCCGCAGAGTAATG 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23358 CTGATGCTCAGGGTATCCGGGACATCTTGATCATCTCCACCCCGCGACCTGCCGATG 23417

QY 181 ATTAATTTCTGAAGACTTTGAGGATAAGCTTGGGATACACATTTACATGCTCCCAAGAG 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23418 TTCCAGGCGCTGCTCGGCGACGTTTCGGCATTCGGCATCAACCTGAGCTATGCCGAACAG 23477

QY 241 ACTGAGCCCTTAGAACCCGCTGCCCTCTTGCTCTAGCAAGGACAAAGCTTGGCGGATGA 300
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23478 CCTTCGCCCAACGGCTTCGGAAGCCTTCATCATCGGCGCGGATTTTCGTCGCAACGAT 23537

QY 301 TCTGGCCAGCATCTTTTGTCTCAACAGTGAT 333
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23538 CCCAGCGCTGATCTTCGGGCAACATCTAT 23570

RESULT 14
US-09-096-942-2
; Sequence 2, Application US/09096942
; Patent No. 6027925
; GENERAL INFORMATION:
; APPLICANT: Pollock, Thomas J
; APPLICANT: Mikolajczak, Marcia
; APPLICANT: Yamazaki, Motohide
; APPLICANT: Thorne, Linda
; APPLICANT: Armentrout, Richard W
; TITLE OF INVENTION: Production of Xanthan Gum by Sphingomonas Bacteria
; FILE REFERENCE: seq list for app'l filed from pro. appl
; CURRENT APPLICATION NUMBER: US/09/096,942
; CURRENT FILING DATE: 1998-06-12
; EARLIER APPLICATION NUMBER: 60/049,428
; EARLIER FILING DATE: 1997-06-12
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 28804
; TYPE: DNA
; ORGANISM: Sphingomonas sp. S88
```

```
US-09-096-942-2

Query Match          3.7%; Score 40.2; DB 3; Length 28804;
Best Local Similarity 45.0%; Pred. No. 0.033;
Matches 150; Conservative 0; Mismatches 183; Indels 0; Gaps 0;

QY 1 ATGAAGCCCTCATCTTGTGCGGGGTTTCGGAACCCGCTTCGGCCCTTTGACTCTGAGC 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23238 ATGAAGGTATCATCTTTCGGGGGCGAGCGGCGCTCTACCCGCAACGCTGTCG 23297

QY 61 TTCCGAAACCCCTCGTGGATTTCGAAACAGCCCATGATCTGCACCAAGATCGAAGCT 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23298 ATCTCGAACAGCTGCTTCCCGTCTATGACAAGCCCATGATCTTACCCCTGTCGGTG 23357

QY 121 TTGAAAGAAAGTTGGGTCACAGAGTGGTTTGGCTATCAACTATGCCCGCAGAGTAATG 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23358 CTGATGCTCAGGGTATCCGGGACATCTTGATCATCTCCACCCCGCGACCTGCCGATG 23417

QY 181 ATTAATTTCTGAAGACTTTGAGGATAAGCTTGGGATACACATTTACATGCTCCCAAGAG 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23418 TTCCAGGCGCTGCTCGGCGACGTTTCGGCATTCGGCATCAACCTGAGCTATGCCGAACAG 23477

QY 241 ACTGAGCCCTTAGAACCCGCTGCCCTCTTGCTCTAGCAAGGACAAAGCTTGGCGGATGA 300
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23478 CCTTCGCCCAACGGCTTCGGAAGCCTTCATCATCGGCGCGGATTTTCGTCGCAACGAT 23537

QY 301 TCTGGCCAGCATCTTTTGTCTCAACAGTGAT 333
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23538 CCCAGCGCTGATCTTCGGGCAACATCTAT 23570

RESULT 15
US-09-096-867-2
; Sequence 2, Application US/09096867
; Patent No. 6030817
; GENERAL INFORMATION:
; APPLICANT: Pollock, Thomas J
; APPLICANT: Mikolajczak, Marcia
; APPLICANT: Yamazaki, Motohide
; APPLICANT: Thorne, Linda
; APPLICANT: Armentrout, Richard W
; TITLE OF INVENTION: Production of Xanthan Gum by Sphingomonas Bacteria
; FILE REFERENCE: seq list for app'l filed from pro. appl
; CURRENT APPLICATION NUMBER: US/09/096,867
; CURRENT FILING DATE: 1998-06-11
; EARLIER APPLICATION NUMBER: 60/049,428
; EARLIER FILING DATE: 1997-06-12
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 28804
; TYPE: DNA
; ORGANISM: Sphingomonas sp. S88
US-09-096-867-2

Query Match          3.7%; Score 40.2; DB 3; Length 28804;
Best Local Similarity 45.0%; Pred. No. 0.033;
Matches 150; Conservative 0; Mismatches 183; Indels 0; Gaps 0;

QY 1 ATGAAGCCCTCATCTTGTGCGGGGTTTCGGAACCCGCTTCGGCCCTTTGACTCTGAGC 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23238 ATGAAGGTATCATCTTTCGGGGGCGAGCGGCGCTCTACCCGCAACGCTGTCG 23297

QY 61 TTCCGAAACCCCTCGTGGATTTCGAAACAGCCCATGATCTGCACCAAGATCGAAGCT 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23298 ATCTCGAACAGCTGCTTCCCGTCTATGACAAGCCCATGATCTTACCCCTGTCGGTG 23357

QY 121 TTGAAAGAAAGTTGGGTCACAGAGTGGTTTGGCTATCAACTATGCCCGCAGAGTAATG 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23358 CTGATGCTCAGGGTATCCGGGACATCTTGATCATCTCCACCCCGCGACCTGCCGATG 23417

QY 181 ATTAATTTCTGAAGACTTTGAGGATAAGCTTGGGATACACATTTACATGCTCCCAAGAG 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23418 TTCCAGGCGCTGCTCGGCGACGTTTCGGCATTCGGCATCAACCTGAGCTATGCCGAACAG 23477

QY 241 ACTGAGCCCTTAGAACCCGCTGCCCTCTTGCTCTAGCAAGGACAAAGCTTGGCGGATGA 300
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23478 CCTTCGCCCAACGGCTTCGGAAGCCTTCATCATCGGCGCGGATTTTCGTCGCAACGAT 23537

QY 301 TCTGGCCAGCATCTTTTGTCTCAACAGTGAT 333
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23538 CCCAGCGCTGATCTTCGGGCAACATCTAT 23570
```

Db 23418 TTCCAGCGCTGCTCGGCACGGTTTCGGGATTCGGGCATCAACCTGAGCTATGCCGAACAG 23477
OY 241 ACTGAGCCCTTAGGACCGCTGGCCCTCTGCTCTAGCAAGGACAAAGCTTCGCCATGCA 300
Db 23478 CCTCGCCCAACGGCCTTCGGCAAGCCTTCATCATCGGCCCGGATTCGTCGGCAGGAT 23537
OY 301 TCTGGCCAGCCATTCTTTGTCTCTACACAGTGAT 333
DD 23538 CCCAGCGCGTGTATCTCGGCGACAAACATCTAT 23570

Search completed: November 27, 2002, 04:07:04
Job time : 185 secs

Result No.	Score	Query Match	Length	DB	ID	Description
1	187	17.2	212	10	US-09-294-093B-2164	Sequence 2164, App
2	151.4	13.9	364	10	US-09-878-574-2725	Sequence 2725, App
3	130.6	12.0	269	10	US-09-878-574-8673	Sequence 8673, App
4	79.2	7.3	272	10	US-09-878-574-10462	Sequence 10462, A
c 5	73.2	6.7	787	10	US-09-954-773A-20	Sequence 20, Appl
6	55.2	5.1	282	10	US-09-294-093B-3871	Sequence 3871, Ap
7	46.2	4.3	678	10	US-09-738-973-155	Sequence 155, App
c 8	38	3.5	167343	10	US-09-962-436-281	Sequence 281, App
9	38	3.5	167343	10	US-09-964-824-273	Sequence 273, App
c 10	35.2	3.2	808	10	US-09-954-773A-19	Sequence 19, Appl
11	35.2	3.2	879	9	US-09-860-846-11	Sequence 11, Appl
12	35.2	3.2	879	10	US-09-861-289-11	Sequence 11, Appl
c 13	35.2	3.2	13613	9	US-09-860-846-3	Sequence 3, Appl
c 14	35.2	3.2	13613	10	US-09-861-289-3	Sequence 3, Appl
15	34.8	3.2	447	10	US-09-770-444-767	Sequence 767, App
16	34.6	3.2	396	10	US-09-878-574-757	Sequence 757, App
c 17	34.4	3.2	1130	10	US-09-939-980-248	Sequence 248, App
18	33.8	3.1	410	10	US-09-878-574-380	Sequence 380, App
c 19	33.6	3.1	13808	10	US-09-070-927A-271	Sequence 271, App

DD 122 ACATTACNGGC-TGGGTCTTTATCTAGACTCGATTAGGAAGAAATCAGCTGCCAAGCTAG 180
QY 737 CTACTGACACATGTTTGGCAATGTGCTG 768
DB 181 CTACTGACACATGTTTGGCAATGTGCTG 212

RESULT 2

US-09-878-574-2725
; Sequence 2725, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 2725
; LENGTH: 364
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: LIB3028-021-Q1-B1-C7
US-09-878-574-2725

Query Match 13.9%; Score 151.4; DB 10; Length 364;
Best Local Similarity 74.3%; Pred. No. 1.5e-39;
Matches 191; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
QY 1 ATGAGGCCCTCATCTTCTCGGGGGTTTCGGAACCCGCTTCGGCCTTTGACTCTGAGC 60
DB 108 ATGAAGGCAATTTCTTGTGGGGATTGGAACAAGCTGAGGCCACTGACACTCACT 167
QY 61 TTCCCGAAACCCCTCGTGGATTTTGCACAAAGCCCATGATTCGACACAGATCGAAGCT 120
DB 168 TTCCCTAAGCCCTCTGTTGATTTGCTAACAGCCCTATGATTTTGATCAGATAGAGCC 227
QY 121 TTGAAGAAGATTTGGGTACAGAGGTGTTTGGCTATCAACTATATGCCCCAGAGTAATG 180
DB 228 CTTAAGGCCATTTGGAGTCTGAGTGTGCTAGCCATCAATTAACCAACAGAGTTATG 287
QY 181 ATTAATTTCTTGAAGACTTTTGAAGATAGCTTGGCATCACAAATTAATGCTCCCAAGAG 240
DB 288 TTGAATTTCTTGAAGATTTTGAATCAAGCTCGGCATCAAGATCACATGTTCTCAGGA 347
QY 241 ACTGAGCCCTTAGGAC 257
DB 348 ACTGAACCACTGGGAAC 364

RESULT 3

US-09-878-574-9673
; Sequence 8673, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 8673
; LENGTH: 269

TVDE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 701101510H1
US-09-878-574-8673
Query Match 12.0%; Score 130.6; DB 10; Length 269;
Best Local Similarity 70.3%; Pred. No. 7.9e-33;
Matches 189; Conservative 0; Mismatches 79; Indels 1; Gaps 1;
QY 230 GGTCCCAAGACACTGACCCCTTAGGAACCGCTGGCCCTCTTGTCTTAGCAAGGACAAGC 209
DB 1 GTTCTCAGGAACCTGAAACCATTTGGGAACAGCAGGCTCTCTGGCTCTTCCAGGATAAGC 60
QY 290 TTGGGATGATCTGGCC-AGCCATTTCTTGTCTCTCAACACTGATGTCATAGCGAATAC 348
DB 61 TGATACATCTCTCTGGAGAACCTTTTGTTCATCAACAGTATGTTATCAGTAGAT 120
QY 349 CCATTTGCTGAACCTCATCAAAATTCACAAGTGTCTATGGTGTGAGGCAACAATATGCTC 408
DB 121 CCACTCAAGAAATGATTAATTCATACAAACCATGAGGAGGCTTCCATATATGTA 180
QY 409 ACTAAGGTGATGAACCATCAAAATACGGTGTGTTGTTATGAGGAGGCAACTGGCAGG 468
DB 181 ACCAAGTTGACGAGCCATCAAGTACGGCTGCTGTGATGGAAGACACACAGGCCAG 240
QY 469 GTGGAAGGTTTGTGAGAGCCAAAAT 497
DB 241 GTTGATAAATTTGTTGAACCAACCAAAAT 269

RESULT 4
US-09-878-574-10462
; Sequence 10462, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 10462
; LENGTH: 272
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 700966702H1
US-09-878-574-10462

Query Match 7.3%; Score 79.2; DB 10; Length 272;
Best Local Similarity 72.9%; Pred. No. 5.5e-16;
Matches 102; Conservative 0; Mismatches 38; Indels 0; Gaps 0;
QY 1 ATGAAGCCCTCATCTTCTGCGGGGTTTCGGAACCCGCTTCGGCCTTTGACTCTGAGC 60
DB 132 ATGAAGCATTTGATCTGTGTTGGGGGTTTGGAAACAAGGCTGAGGCCACTCAGCTCAGT 191
QY 61 TTCCCGAAACCCCTCGTGGATTTTGCACAAAGCCCATGATTTGTCACACAGATCGAAGCT 120
DB 192 TTCTTAGGCTCTGTTGATTTTGCACAAACCAACCTATGATTCGTCATCAGATAGAGCC 251
QY 121 TTGAAGAAGCTTGGGTCAC 140
DB 252 CTTAAGGCCCATTTGAGTTAC 271

RESULT 5
us-09-954-773A-20/c
; Sequence 20, Application US/09954773A

Patent No. US20020129402A1
GENERAL INFORMATION:
APPLICANT: Southern Illinois University
APPLICANT: Lightfoot, David A
APPLICANT: Gibson, Paul T
APPLICANT: Meksem, Khalid
TITLE OF INVENTION: SOYBEAN SUDDEN DEATH SYNDROME RESISTANT SOYBEANS, SOYBEAN CYST NEMATODE RESISTANT SOYBEANS AND METHODS OF BREEDING AND IDENTIFYING RESISTANT SOYBEANS
FILE REFERENCE: 1268/2/2
CURRENT APPLICATION NUMBER: US/09/954,773A
PRIORITY FILING DATE: 2002-01-22
PRIORITY FILING DATE: 1998-01-14
PRIORITY FILING DATE: 1998-01-14
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn version 3.1
SEQ ID NO 20
LENGTH: 787
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1)..(787)
OTHER INFORMATION: n is a, c, g, or t/u
US-09-954-773A-20

Query Match 6.7%; Score 73.2; DB 10; Length 787;
Best Local Similarity 56.0%; Pred. No. 1.1e-13;
Matches 227; Conservative 0; Mismatches 151; Indels 27; Gaps 5;
QY 524 GGATTACTTACTGAACCATCTGTCTTGGACGCGATTGAGCTGAGCGCCCAACATCAATG 583
DB 563 GGGTAACTGGATGAACCTTGGTATGGANAGATNGAGNGACACCATTTCTTGGGA 504
QY 584 AGAAGAGCTCTCCCTCAAAATGACGTCATCAACAGCTCTATGCAATGGTCTTCC-- 641
DB 503 AGACAGAGGGTTTCCAAATATGTTTCNGATAGAAGCTGNATGCAATGGTCTTCCAC 444
QY 642 ----AGTTTGGATGATGTTGTCAGCTAGGAGTACATCTACTGCGTCTTT 697
DB 443 AGGAAGATTNAGANGAANGGACAA--AGGACTATATTTNNGANTGAGACTCN 387
QY 698 ATCTAGCTGATTAGGAAGA---AATCAGCTGCCAGCTAGCTAGGAGCACATGTTG 754
DB 386 ACNTGAACCTGGGAGGAATAATNTCTTAGTGAACNGGCAATGACNTCANACATANTG 327
QY 755 TTGGCAATGCTGTTGATGATGAGCGCCCAAGATTGGAGAAGTTGCTGATTGTCCTG 814
DB 326 TGGGAAATGCTGTTGATGATGAGCTGCCAA-ATNGGTGATGATGTCAAATTTGGTCC-- 270
QY 815 ATGTGCCATTGGACCTGGTGTGTTGTGGAGGACGCGTGGAGCTTCCCGCTGCACTG 874
DB 269 -----ATCNGGATGATCATTTGAGAAGAGTTGGGCTCAAACTNACACTG 222
QY 875 TCATCGCGCGCTGGCTATCAAGAAGCATGCTTGCATCTCAACA 919
DB 221 TTATGTGAAGCACTCGTGTGTTAAAGCAATTTCTTATATCNGACA 177

RESULT 6
US-09-294-093B-3871
Sequence 3871, Application US/09294093B
Patent No. US20010051335A1
GENERAL INFORMATION:
APPLICANT: Lalgudi, Raghunath, V.
APPLICANT: Ito, Laura, Y.
APPLICANT: Sherman, Bradley, K.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL
FILE REFERENCE: PL-0009 US
CURRENT APPLICATION NUMBER: US/09/294,093B
PRIORITY FILING DATE: 1999-04-16
PRIORITY FILING DATE: 1999-04-16
PRIORITY FILING DATE: April 21, 1998
NUMBER OF SEQ ID NOS: 6207

SOFTWARE: PERL Program
SEQ ID NO 3871
LENGTH: 282
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID No. US20010051335A1 700352730H1
NAME/KEY: unsure
LOCATION: 50, 185, 218
OTHER INFORMATION: a, t, c, g, or other
US-09-294-093B-3871
Query Match 5.1%; Score 55.2; DB 10; Length 282;
Best Local Similarity 89.9%; Pred. No. 4.1e-08;
Matches 71; Conservative 0; Mismatches 4; Indels 4; Gaps 1;
QY 1 ATGAAGGCCCTCATTTCTGTCGGGGTTTCGGAACCCGCTTCGGCCTTTGACTGTGAGC 60
DB 208 ATGAAGGCCCNCAATTTCTGTCGGGGTTTCGGAACCCGCTTCGGCCTTAGATT----GG 263
QY 61 TTCCCGGAACCCCTCGTGG 79
DB 264 TTCCCGGAACCCCTCGTGG 282

RESULT 7
US-09-738-973-155
Sequence 155, Application US/09738973
Patent No. US20020110563A1
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Henderson, Robert A.
APPLICANT: Lodes, Michael J.
APPLICANT: Fling, Steven P.
APPLICANT: Mohamath, Raodoh
APPLICANT: Alghate, Paul A.
APPLICANT: Secrist, Heather
APPLICANT: Indrias, Carol Yoseph
APPLICANT: Benson, Darin R.
APPLICANT: Elliot, Mark
APPLICANT: Mannion, Jane
APPLICANT: Kalos, Michael D.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS OF LUNG CANCER
TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 21021.475C9
CURRENT APPLICATION NUMBER: US/09/738,973
CURRENT FILING DATE: 2000-12-14
NUMBER OF SEQ ID NOS: 587
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 155
LENGTH: 678
TYPE: DNA
ORGANISM: Homo sapien
US-09-738-973-155
Query Match 4.3%; Score 46.2; DB 10; Length 678;
Best Local Similarity 49.5%; Pred. No. 6.7e-05;
Matches 151; Conservative 0; Mismatches 148; Indels 6; Gaps 1;
QY 235 CAAGAGACTGAGCCCTTAGGAACGCTGGCCCTCTTGTCTAGCAAGGACCAAGCTTGG 294
DB 350 CAGGAATTTGCCCCCTTAGGCACAGGGGTGCTCTTACCATTTCGAGACAGATCTCG 409
QY 295 GATGGATCTGGCCAGCCATCTTTGTCTCAACAGTGTATGTCATAAGCGAATACCCATT 354
DB 410 GCTGGGAGCCCGAGGATCTTCTGCTCAATGCTGATGCTGCTCGACTTCCCTTG 469
QY 355 GCTGAATCATCAAAATTTCACAAGTGTATGTTGGTGGAGCAACAATTTATGTTCACTAAG 414
DB 470 AGTGTATGTTGGAAGCCACCGACGCGCTACCCCTTTCTTACTCTTGGCACTAGG 529
QY 415 GTGGATG-----AACCATCAAAATACGGTGTGTTGGTATGAGGAGGCAACTGGCAGG 468

Db 530 GCTAACAGGACGCAATCCCTCACTACCTGCGTGCATGTTGAGATCCACAGACACACGAG 589
QY 469 GTGGAAAGGTTTGTGAGAGCCCAAAATATTCTGGGTAACAAAGATCAATGCTGGGATT 528
Db 590 GTATTGACATGCTGGAGAAACCCACACATTATCATGTCACATCACTACCTGGGCAAC 649
QY 529 TACTTT 533
Db 650 TACCT 654

RESULT 8

US-09-962-436-281/c
; Sequence 281, Application US/09962436
; Patent No. US20020081301a1
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689290-75
; CURRENT APPLICATION NUMBER: US/09/962,436
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/60/235,082
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/234,924
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 568
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 281
; LENGTH: 167343
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-436-281

Query Match 3.5%; Score 38; DB 10; Length 167343;
Best Local Similarity 52.5%; Pred. No. 1.4;
Matches 83; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 929 GCTGGCAGCTCACTGTGGTCAATGGCGGATAGAGATATGACTATCTCTGGGGAGG 988
Db 120438 GCCAGGATTCAGTGTAAAGGATGGCCAGGAGAGGATGTTGGGATTGGGAACGAG 120379
QY 989 ATGTCATGTGTGATGAGGTAGAGCAATGGCGGTGTGTTCTCCACATAAAGAGA 1048
Db 120378 AATGGGATGTAGCTGTCAAACTCTAGAGGATCAGTCTGATAGATCCTACAATAAAGCCT 120319
QY 1049 TCAAGTCAGCACTCTGAAGCCTGAGATCGTCATGTGA 1086
Db 120318 TTAAGCTTGGAACTGTTGAATGTGAAGGTAAGAGA 120281

RESULT 9

US-09-964-824A-273/c
; Sequence 273, Application US/09964824A
; Patent No. US20020102531a1
; GENERAL INFORMATION:
; APPLICANT: Horrigan, Stephen
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689290-73
; CURRENT APPLICATION NUMBER: US/09/964,824A
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/60/236,033
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,032
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,028
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 583
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 273
; LENGTH: 167343

; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-964-824A-273
Query Match 3.5%; Score 38; DB 10; Length 167343;
Best Local Similarity 52.5%; Pred. No. 1.4;
Matches 83; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
QY 929 GCTGGCAGCTCACTGTGGTCAATGGCGGATAGAGATATGACTATCTCTGGGGAGG 988
Db 120438 GCCAGGATTCAGTGTAAAGGATGGCCAGGAGGAGGATGTTGGGAAACGAG 120379
QY 989 ATGTCATGTGTGATGAGGTAGAGCAATGGCGGTGTGTTCTCCACATAAAGAGA 1048
Db 120378 AATGGGATGTAGCTGTCAAACTCTAGAGGATCAGTCTGATAGATCCTACAATAAAGCCT 120319
QY 1049 TCAAGTCAGCACTCTGAAGCCTGAGATCGTCATGTGA 1086
Db 120318 TTAAGCTTGGAACTGTTGAATGTGAAGGTAAGAGA 120281

RESULT 10

US-09-954-773A-19/c
; Sequence 19, Application US/09954773A
; Patent No. US20020129402a1
; GENERAL INFORMATION:
; APPLICANT: Southern Illinois University
; APPLICANT: Lightfoot, David A
; APPLICANT: Gidson, Paul T
; APPLICANT: Meksem, Khalid
; TITLE OF INVENTION: SOYBEAN SUDDEN DEATH SYNDROME RESISTANT SOYBEANS, SOYBEAN CYST
; FILE REFERENCE: 1268/2/2
; CURRENT APPLICATION NUMBER: US/09/954,773A
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: 09/007,119
; PRIOR FILING DATE: 1998-01-14
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 808
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(808)
; OTHER INFORMATION: n is a, c, g, or t/u
US-09-954-773A-19

Query Match 3.2%; Score 35.2; DB 10; Length 808;
Best Local Similarity 46.7%; Pred. No. 0.3;
Matches 196; Conservative 0; Mismatches 201; Indels 23; Gaps 4;

QY 518 ATGCTGGGATTACTTACTGAACCCATCTGTCTTGACCCGATT-GAGCTGAGGCCACA 576
Db 707 ATNCGGGTTAAAGTGGAGGAACCCCTTNGGAATGGANAGAAATGGAGTGAACCCAAAT 648
QY 577 TCAATTCAGAAAGAGCTTCCCTCAAATTCACGCTGATCAACAGCTCTATCAATGTC 636
Db 647 TCTTNGAAACAGAGGGTTTCCNAAAAAATNGTTTGGGAAGAANGCTGGAAGCAAAAGTT 580
QY 637 CTTCCAGGTTTTTGA-----TGGATGTGTGTCAGCTAGGAGCTACATTAAGCTTCCG 692
Db 587 CTTTCNAACAGNANGAATGAGATGGAAATGGCAACGGGATATATTTTNNNGAATGAG 528
QY 693 TCTTTATCTAGACTCGATTAGGAAGAAATCAGCTGCCAAGCTAGCTACTGAGAGCACA--- 749
Db 527 ANTCAACCTGAANNNGGGGAGCAATATTTNTAAGGAACNGGCCAANGACCTTCAAAC 468
QY 750 TCTTTGTTGGCAATGCTGCTGTCATGAGAGCCGCAAGATTGGAGAGAGTTGCTCTGATTGG 809
Db 467 ATACTGNGGGAATGTTGGTGGTGCATGAGACTNCCAAATNGTGGAGNNATGTCAANNNGT 409

QY 810 TCCTGATGTCGCCATTGACCTGGGTGTTGTGAGAGACGGCGTGAGGCTTTCCCGCTG 869
Db 407 TCAT-----CCTGNATGATCATTTGAGAAGGAGNTGGGCTCAAATCTAA 363
QY 870 CACTGTATCGCGCGGTGCTATCAAGAAGCATGCTTGCATCTCAAAACAGCATTTATCGG 929
Db 362 CANTGTNAGGGGAAGCACTCGTGTAAAGACATTTNGTTATATCNGACAATGNGATTNG 303

RESULT 11

US-09-860-846-11
; Sequence 11, Application US/09860846
; Patent No. US20020164742A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/860.846
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 879
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-860-846-11

Query Match 3.2%; Score 35.2; DB 9; Length 879;
Best Local Similarity 44.2%; Pred. No. 0.32;
Matches 145; Conservative 0; Mismatches 183; Indels 0; Gaps 0;

QY 1 ATGAAGCCCTCATCTCTGTGCGGGGTTTCGGAACCCCGCTTCGGCCCTTTGACTCTGAGC 60
Db 1 ATGAAGGAATAGTCTCGCGCGGGAGCGAACTCGCGTGCATCCGCGACCTCGGTC 60
QY 61 TTCCGAAACCCCTCGTGGATTTCGAAACAGCCCATGATTCGCACCAATCGAAAGCT 120
Db 61 ATTTGGAAGCATGTTCTCCGGTCTACAACAAACCGATGATCTACTATCCGCTGTCGGTT 120
QY 121 TTGAAGAAGTTGGGTCACAGAGTGTGTTTGGCTATCACTATCGCCCGAGAGTAATG 180
Db 121 CTCATGCTCGCGGTATTCGGAGATTCAAAATCATCTCGACCCCGCAGCATCGAACTC 180
QY 181 ATTAATTTCTGAAGGACTTTGAGGATAAGCTTGGCAATGATTCGATGCTCCCAAGAG 240
Db 181 TTCCAGTCGCTTCTCGGAACGGCAGGCACCTGGGAATAGAACTGATCGGTCCAG 240
QY 241 ACTGAGCCCTTAGAAGCGTGGCCCTCTGCTCTAGCAAGGACAAAGCTTGGGGATGGA 300
Db 241 AAAGAGCCCGAGGAATCGCGGCGGCACTTCTCGTCGGAGCGGAGCGAGCATCGGCGAC 300
QY 301 TCTGCGCAGCATCTTGTCTCTCAACA 328
Db 301 ACCTGCGCCTGTATCTCGGGGCAACA 328

RESULT 12

US-09-861-289-11
; Sequence 11, Application US/09861289
; Patent No. US20020110897A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/861.289

; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 879
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-861-289-11

Query Match 3.2%; Score 35.2; DB 10; Length 879;
Best Local Similarity 44.2%; Pred. No. 0.32;
Matches 145; Conservative 0; Mismatches 183; Indels 0; Gaps 0;

QY 1 ATGAAGCCCTCATCTCTGTGCGGGGTTTCGGAACCCCGCTTCGGCCCTTTGACTCTGAGC 60
Db 1 ATGAAGGAATAGTCTCGCGCGGGAGCGAACTCGGCTGCATCGGGGAGCTCGGTC 60
QY 61 TTCCGAAACCCCTCGTGGATTTCGAAACAGCCCATGATTCGCACCAATCGAAAGCT 120
Db 61 ATTTGGAAGCATGTTCTCCGGTCTACAACAAACCGATGATCTACTATCCGCTGTCGGTT 120
QY 121 TTGAAGAAGTTGGGTCACAGAGTGTGTTTGGCTATCACTATCGCCCGAGAGTAATG 180
Db 121 CTCATGCTCGCGGTATTCGGAGATTCAAAATCATCTCGACCCCGCAGCATCGAACTC 180
QY 181 ATTAATTTCTGAAGGACTTTGAGGATAAGCTTGGCAATGATTCGATGCTCCCAAGAG 240
Db 181 TTCCAGTCGCTTCTCGGAACGGCAGGCACCTGGGAATAGAACTGATCGGTCCAG 240
QY 241 ACTGAGCCCTTAGAAGCCGCTGGCCCTCTGCTCTAGCAAGGACAAAGCTTGGGGATGGA 300
Db 241 AAAGAGCCCGAGGAATCGCGGCGGCACTTCTCGTCGGAGCGGAGCGACATCGGCGAGC 300
QY 301 TCTGCGCAGCATCTTGTCTCTCAACA 328
Db 301 ACCTGCGCCTGTATCTCGGGGCAACA 328

RESULT 13

US-09-860-846-3/c
; Sequence 3, Application US/09860846
; Patent No. US20020164742A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/860.846
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 13613
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-860-846-3

Query Match 3.2%; Score 35.2; DB 9; Length 13613;
Best Local Similarity 44.2%; Pred. No. 2.1;
Matches 145; Conservative 0; Mismatches 183; Indels 0; Gaps 0;

QY 1 ATGAAGCCCTCATCTCTGTGCGGGGTTTCGGAACCCCGCTTCGGCCCTTTGACTCTGAGC 60
Db 9989 ATGAAGGAATAGTCTCGCGCGGGAGCGAACTCGGCTGCATCGGGGAGCTCGGTC 9930
QY 61 TTCCGAAACCCCTCGTGGATTTCGAAACAGCCCATGATTCGCAACAGATCGAAAGCT 120

Db 9929 ATTTCGAGCAGATCTTCGGGTCTACACAAACCCGATGATCTACTATCCGCTGTTCGGTT 9870
QY 121 TTGAAAGAGTTGGGTACAGAGCGTGTTCGGCTATCAACTATCGCCACAGAGTAATG 180
Db 9869 CTCATGCTCGGCGGTATTCGCGAGATTCAAATCATCTCGACCCGCCAGCACATCGAACTC 9810
QY 181 ATTAATTTCTTGAAGACATTTGAGGATATAGCTTGGATACAAATTACATGCTCCCAAGAG 240
Db 9809 TTCCAGTCGCTTCTCGAAACGCCAGCGACCTGGGTAATAGAACTCGACTATCGGTCAG 9750
QY 241 ACTGAGCCCTTAGGAACCGCTGGCCCTCTTGTCTTAGCAAGGACAAAGCTTGGCGGATGCA 300
Db 9749 AAGAGCCCGGAGATTCGGGACGACATCTCGTCGGAGCCGAGCACATCGCGGACGAC 9690
QY 301 TCTGGCCAGCCATTCCTTGTCTCTCAACA 328
Db 9689 ACCTGGCCCTGATCTCTGGCGACACAA 9662

RESULT 14

US-09-861-289-3/c
; Sequence 3, Application US/09861289
; Patent No. US20020110897A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/861,289
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 13613
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-861-289-3

Query Match 3.2%; Score 95.2; DB 10; Length 13613;
Best Local Similarity 44.2%; Pred. No. 2.1;
Matches 145; Conservative 0; Mismatches 183; Indels 0; Gaps 0;

QY 1 ATGAGAGCCCTCATCTTCTCGGGGTTCGGAACCCGCTTCGGCTTTGACTCTTGAGC 60
Db 9989 ATCAAGCGAATATCTCTGCGCGCGGAGCGGAACCTCGCTGCATCCGGCGACCTCGGTC 9930
QY 61 TTCCGAGACCCCTCTGTGATTTTGCACAAACCCCATCATCTTGCACACAGATCGAAGCT 120
Db 9929 ATTTGAGCAGATCTTCGGGTCTACAAACACCGATATCTACTATCCGCTGTTCGGTT 9870
QY 121 TTGAAAGAGTTGGGTACAGAGTGTTCGGCTATCAACTATCGCCACAGAGTAATG 180
Db 9869 CTCATGCTCGGCGGTATTCGCGAGATTCAAATCATCTCCACCCCGACACATCGAACTC 9810
QY 181 ATTAATTTCTTGAAGACATTTGAGGATAAGCTTGGCATACAAATTACATGCTCCCAAGAG 240
Db 9809 TTCCAGTCGCTTCTCGAAACGCCAGCGACCTGGGTAATAGAACTCGACTATCGGTCAG 9750
QY 241 ACTGAGCCCTTAGGACCGCTGGCCCTCTTGTCTTAGCAAGGACAAAGCTTGGCGGATGCA 300
Db 9749 AAGAGCCCGGAGATTCGGGACGACATCTCGTCGGAGCCGAGCACATCGGCGACGAC 9690
QY 301 TCTGGCCAGCCATTCCTTGTCTCTCAACA 328
Db 9689 ACCTGGCCCTGATCTCTGGCGACACAA 9662

RESULT 15

US-09-770-444-767

; Sequence 767, Application US/09770444
; Patent No. US20020023280A1
; GENERAL INFORMATION:
; APPLICANT: Gotlach, Jorn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Krieker, Maja
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hurban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; TITLE OF INVENTION: thaliana
; FILE REFERENCE: 2027 (PARA-016PRV)
; CURRENT APPLICATION NUMBER: US/09/770,444
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,502
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 767
; LENGTH: 447
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-770-444-767

Query Match 3.2%; Score 34.8; DB 10; Length 447;
Best Local Similarity 58.8%; Pred. No. 0.27;
Matches 60; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 31 CGAACCCGCTTCGGCTTTGACTCTGAGCTTCGGAACCCCTCGTGGATTTGCAAC 90
Db 135 GGCACTCGATTCGGCCATTTCGCTGAATATTCGAAAGCTCTGTTCTCTATTGCGGA 194
QY 91 AAGCCCATGATTCGACACGATCGAGCTTTGAAAGAGTT 132
Db 195 CAACCAATGGTCATCATCCATTTTCAGCTTGTAAGAAAT 236

Search completed: November 27, 2002, 05:08:44
Job time : 353 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 2002, 02:38:36 ; Search time 1981 Seconds
(without alignments)
8878.495 Million cell updates/sec

Title: US-09-374-967-1

Perfect score: 1086

Sequence: 1 atgaaggccctcattcttgt.....agcctgagatgctatgtga 1086

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estopl:*
7: em_estro:*
8: em_esthc:*
9: gb_est1:*
10: gb_est2:*
11: gb_est3:*
12: gb_est4:*
13: gb_est5:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1068.4	98.4	1542	11	AY108204 Zea mays
2	753.8	69.4	980	12	BG321249 Zm04_04b0
3	538.2	49.6	694	12	BG049329 OVL18_E1
4	490.4	45.2	836	10	BE455732 HVSMEg001
5	486	44.8	524	13	BM324237 PIC1_26_A
6	483.4	44.5	597	10	BE405374 WHE1213_F

c	7	478.6	44.1	686	13	BJ320079	BJ320079
c	8	476.4	43.9	802	12	BG582583	BG582583
c	9	460.2	42.4	697	13	BJ318153	BJ318153
c	10	459.4	42.3	654	13	BI075102	BI075102
c	11	447.4	41.2	651	10	AV935887	AV935887
c	12	442.8	40.8	794	12	BE844694	BE844694
c	13	438	40.3	619	14	BQ806351	BQ806351
c	14	437.4	40.3	673	10	AV928662	AV928662
c	15	435	40.1	808	12	BG414225	BG414225
c	16	434.4	40.0	774	13	BM413164	BM413164
c	17	433	39.9	789	13	BI421941	BI421941
c	18	431.6	39.7	474	13	BM499735	BM499735
c	19	427	39.3	612	10	AV925892	AV925892
c	20	421.4	38.8	794	12	BG351126	BG351126
c	21	420	38.7	684	14	BQ970876	BQ970876
c	22	417.4	38.4	870	12	BF254904	BF254904
c	23	415.2	38.2	609	10	AV933434	AV933434
c	24	414.4	38.2	698	17	BH875336	BH875336
c	25	407.2	37.5	642	13	BQ265250	BQ265250
c	26	396.4	36.5	687	14	BQ506346	BQ506346
c	27	396	36.5	665	13	BJ314593	BJ314593
c	28	391	36.0	623	14	BM779448	BM779448
c	29	390.6	36.0	650	10	AM692513	AM692513
c	30	387.6	35.7	618	14	BM779156	BM779156
c	31	387.4	35.7	706	11	AY105397	AY105397
c	32	385.6	35.5	658	14	BQ789419	BQ789419
c	33	384	35.4	674	12	BG587608	BG587608
c	34	381.8	35.2	651	10	AM287555	AM287555
c	35	381	35.1	669	14	BQ487033	BQ487033
c	36	381	35.1	729	13	BI933297	BI933297
c	37	378.6	34.9	694	14	BQ743338	BQ743338
c	38	378.2	34.8	774	13	BI932250	BI932250
c	39	377.6	34.8	639	13	BJ259364	BJ259364
c	40	372.6	34.3	679	12	BG350520	BG350520
c	41	369	34.0	584	13	BM328918	BM328918
c	42	368.6	33.9	585	13	BM328644	BM328644
c	43	367.4	33.8	637	10	AM623911	AM623911
c	44	366	33.7	601	10	BE204724	BE204724
c	45	364.2	33.5	585	13	BM499734	BM499734

ALIGNMENTS

RESULT 1
AY108204
LOCUS AY108204 1542 bp mRNA linear HTC 25-MAY-2002
DEFINITION Zea mays PCO075514 mRNA sequence.
ACCESSION AY108204
VERSION AY108204.1 GI:21211282
KEYWORDS HTC.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 1542)
AUTHORS Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
TITLE Zea Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
JOURNAL Unpublished (2002)
REFERENCE 2 (bases 1 to 1542)
AUTHORS Coe,E.C.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
FEATURES
source
1..1542
/organism="Zea mays"
/db_xref="MaizeDB:634165"
/db_xref="taxon:4577"
/clone="PCO075514"

```
/clone_lib="Maize Mapping Project/DuPont Cornsensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed Dupont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"
BASE COUNT      377 a 341 c 408 g 416 t
ORIGIN
Query Match      98.4%; Score 1068.4; DB 11; Length 1542;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1075; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 1 ATGAAGGCCCTCATCTTCTGGGGGTTTCGAAACCCGCTTCGGCCCTTGAAGCTGAGC 60
DB 202 ATGAAGGCCCTCATCTTCTGGGGGTTTCGAAACCCGCTTCGGCCCTTGAAGCTGAGC 261
QY 61 TTCCCGAAACCCCTGGTGGATTTTGAACAAAGCCCATGTTCTGCACACATCGAAGCT 120
DB 262 TTCCCGAAACCCCTGGTGGATTTTGAACAAAGCCCATGTTCTGCACACATCGAAGCT 321
QY 121 TTGAAGAAGTTGGGTCACAGAGTGTTTGGTATCACTATCGCCCGAGAGTAATG 180
DB 322 TTGAAGAAGTTGGGTCACAGAGTGTTTGGTATCACTATCGCCCGAGAGTAATG 381
QY 181 ATTAATTTCTTTCAGGACTTTTCAGCAATGCTTGGCATTACAAATTTACATGCTCCCAAGG 240
DB 382 ATTAATTTCTTTCAGGACTTTTCAGCAATGCTTGGCATTACAAATTTACATGCTCCCAAGG 441
QY 241 ACTGAGCCCTTAGGAACCCCTGGCCCTCTGCTAGCAAGGCAAGCTTCGGATGGA 300
DB 442 ACTGAGCCCTTAGGAACCCCTGGCCCTCTGCTAGCAAGGCAAGCTTCGGATGGA 501
QY 301 TCTGCGCAGCCATCTTTTGTCTCAACAGTATGTCATAGCGAATATCCCATTTTCTGAA 360
DB 502 TCTGCGCAGCCATCTTTTGTCTCAACAGTATGTCATAGCGAATATCCCATTTTCTGAA 561
QY 361 CTCATCAATTTTCAGACTGTCATGCTGCTGAGGCAACAAATTTGCTACATAGGTGAT 420
DB 562 CTCATCAATTTTCAGACTGTCATGCTGCTGAGGCAACAAATTTGCTACATAGGTGAT 621
QY 421 CAACATCAAAATACGGTGTGCTGTTATGAGGAGGCAATGCGCAGGCTGCAAGCTTT 480
DB 622 GAACCACTAATATGTTGTTGTTATGAGGAGGCAATGCGCAGGCTGCAAGCTTT 681
QY 481 GTTGAGAGGCCAAATAATTTTGGGTAAACAGATCAATGCTGGGATTTACTTACTGAAC 540
DB 682 GTTGAGAGGCCAAATAATTTTGGGTAAACAGATCAATGCTGGGATTTACTTACTGAAC 741
QY 541 CCATCTGTCTTGACCGATTTGAGTGAAGGCAACATCAATTTGAGAAGAGTCTTCCT 600
DB 742 CCATCTGTCTTGACCGATTTGAGTGAAGGCAACATCAATTTGAGAAGAGTCTTCCT 801
QY 601 CAAATTCAGCTGATCAACAGCTGTCATCAATGCTGCTCAGGCTTTTTCAGTGGATGTT 660
DB 802 CAAATTCAGCTGATCAACAGCTGTCATCAATGCTGCTCAGGCTTTTTCAGTGGATGTT 861
QY 661 GGTGAGCCTTAGGACTACATTTACTGCTGCTCTTTATCTAGACTCGATTTAGGAAGAA 720
DB 862 GGTGAGCCTTAGGACTACATTTACTGCTGCTCTTTATCTAGACTCGATTTAGGAAGAA 921
QY 721 TCAGCTGCCAGCTAGCTACTTGGAGCAGATGTTTGGCAATGCTGCTGCTGCTGCTGCTG 780
DB 922 TCAGCTGCCAGCTAGCTACTTGGAGCAGATGTTTGGCAATGCTGCTGCTGCTGCTGCTG 981
QY 781 GCCAAGATTCAGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
DB 982 GCCAAGATTCAGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1041
QY 841 GTGAGGAGCCGCTGAGGCTTTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
```

```
DB 1042 GTGAGGAGCCGCTGAGGCTTTCGCCCTGCACCTGTCTATCGCGGCGCTGCTATCAAGAAG 1101
QY 901 CATGCTTGCATCTCAACAGCATTAATCGCTGCGCTGCGCACTCAACTGCTTGCCTAAATGGGACGG 960
DB 1102 CATGCTTGCATCTCAACAGCATTAATCGCTGCGCTGCGCACTCAACTGCTTGCCTAAATGGGACGG 1161
QY 961 ATAGAGAAATATGACTATCTTGGGGAGGATGTTTCATGTTGTGATGAGGTGTACAGCAAT 1020
DB 1162 ATAGAGAAATATGACTATCTTGGGGAGGATGTTTCATGTTGTGATGAGGTGTACAGCAAT 1221
QY 1021 GCGGCTGTTCTCCACATAAAGAGATCAAGTCAAGCATTTCTGAAGCTTGAGATCGTC 1080
DB 1222 GCGGCTGTTCTCCACATAAAGAGATCAAGTCAAGCATTTCTGAAGCTTGAGATCGTC 1281
QY 1081 ATGTGA 1086
DB 1282 ATGTGA 1287

RESULT 2
LOCUS      BG321249      980 bp      mRNA      linear      EST 27-FEB-2001
DEFINITION Zm04_04b07_R Zm04_AAFC_ECORC_cold_stressed_maize_seedlings Zea mays
ACCESSION BG321249
VERSION    HG321249.1 GI:13150927
KEYWORDS   EST.
SOURCE     Zea mays.
ORGANISM   Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE  1 (bases 1 to 980)
AUTHORS   Singh,J.A., Wakui,K., Couroux,P., De Moors,A., Harris,L.J., Hattori
           J.I., Ouellet,F., Robert,L.S., Sprott,D. and Linker,N.A.
TITLE     Expressed Sequence Tags from Cold-Stressed Maize Seedlings
JOURNAL   Unpublished (2001)
COMMENT   Contact: Singh,J.A.
           Eastern Cereal and Oilseed Research Centre
           Agriculture and Agri-food Canada
           KW Neatby Bldg., Central Experimental Farm, Ottawa, Ontario, K1A
           0C6, Canada
           Tel: (613) 759-1662
           Fax: (613) 759-1701
           Email: singhjadem.agr.ca
FEATURES   Location/Qualifiers
            source          1..980
                        /organism="Zea mays"
                        /cultivar="co328"
                        /db_xref="taxon:4577"
                        /clone="Zm04_04b07"
                        /clone_lib="Zm04_AAFC_ECORC_cold_stressed_maize_seedlings"
                        /tissue_type="leaf, crown"
                        /note="Vector: Bluescript SK-/XhoI-ECORI; Site_1: Eco RI;
                        site_2: Xho I; Lower temperature 50 C / hour from 22 to
                        120C; bring to 50 in 1 hour from 120C. Leave at 50C 2 days
                        , photoperiod 16 hours. Light intensity was 125 uE-1.
                        library prepared by in vivo mass excision from amplified
                        library."
BASE COUNT 237 a 225 c 253 g 240 t 25 others
ORIGIN
Query Match      69.4%; Score 753.8; DB 12; Length 980;
Best Local Similarity 96.5%; Pred. No. 2.4e-222;
Matches 769; Conservative 20; Mismatches 6; Indels 2; Gaps 2;
QY 1 ATGAAGGCCCTCATCTTCTGGGGGTTTCGAAACCCGCTTCGGCCCTTGAAGCTGAGC 60
DB 186 ATGAAGGCCCTCATCTTCTGGGGGTTTCGAAACCCGCTTCGGCCCTTGAAGCTGAGC 245
QY 61 TTCCCGAAACCCCTGGTGGATTTTGAACAAAGCCCATGTTCTGCACACATCGAAGCT 120
DB 246 TTCCCGAAACCCCTGGTGGATTTTGAACAAAGCCCATGTTCTGCACACATCGAAGCT 305
```

QY 121 TTGAAGAAGTTGGGTCACAGAGGTGTTTGGCTATCAACTATCGCCAGAGGTAATG 180
|||||
Db 306 TTGAAGAAGTTGGGTCACAGAGGTGTTTGGCTATCAACTACCGCCAGAGGTAATG 365
QY 181 ATTAATTTCTTGAAGGACTTTGAGGATAAGCTTGGCATCAACATTTACATGCTCCCAAGAG 240
|||||
Db 366 ATTAATTTCTTGAAGGACTTTGAGGATAAGCTTGGCATCAACATTTACATGCTCCCAAGAK 425
QY 241 ACTGAGCCCTTAGGAACCGCTGGCCCTCTTGCTCTAGCAAGGACAAGCTTTCGGGATGA 300
|||||
Db 426 ACTRAGCCCTTAGGACCGCTGGCCCTCTTGCTCTAGCAAGGACAAGCTTTCCKGATGA 485
QY 301 TCTGCGCCAGCCATCTTTTGTCTCAACAGTAGTGATGATCAAGCAATACCCATTTTCTGAA 360
|||||
Db 486 TCTRAGCCAGCCATCTTTTGTCTCAACAGTAGTGATGATCAAGCAATACBCATTTCTGAA 545
QY 361 CTCATCAATTTCAAGTAGTGATGCTGGTGAGGACAATATGCTCACTAAGGTGAT 420
|||||
Db 546 CTCATCAATTTCAAGTAGTGATGCTGGTGAGGACAATATGCTCACTAAGGTGAT 605
QY 421 GAACCATCAAAATACGGTGTGTTGTTATGAGGAGGCAACTGGCAGGTTGGAAGGTTT 480
|||||
Db 606 GAACCATCAAAATACGGTGTGTTGTTATGAGGAGGCAACTGGCAGGTTGGAAGGTTT 665
QY 481 GTTGAGAGCCAAAATATTTTGGGTAAACAGATCAATGCTGGGATTTACTTACTGAAC 540
|||||
Db 666 GTTGAGAGCCAAAATATTTTGGGTAAACAGATCAATGCTGGGATTTACTTACTGAAC 725
QY 541 CCATCTGTCTTGACCGCATTTAGCTGAGGCCAACATCAATTTGAGAAAGAGTCTTCCT 600
|||||
Db 726 CCATCTGTCTTGACCGCATTTAGCTGAGGCCAACATCAATTTGAGAAAGAGTCTTCCT 785
QY 601 CAATTCGAGCTGATCAACAGCTCTATGCAATGGTCTTCAGGTTTGGATGATGTT 660
|||||
Db 786 CAATTCGAGCHGATCAACAGCTCTATGCAATGGTCTTCAGGTTTGGATGATGTT 845
QY 661 GGTGAGCCTAGGGACTACATTTACTTGGCTTGGCTTTTATGACTCGATAGGAAGAAA 720
|||||
Db 846 GGTGAGCCTAGGGACTACATTTACTTGGCTTGGCTTTTATGACTCGATAGGAAGAAA 905
QY 721 TCAGCTGCCAAGCTAGCTAGCTAGGAGCACATGTTGTCATGCTGTCATGAGAGC 780
|||||
Db 906 TCAGCTGCCAAGCTAGCTAGGAGCACATGTTGTCATGCTGTCATGAGAGC 963
QY 781 CCCAAGATTGAGAGG 797
|||||
Db 964 CCCAAGATTGAGAGG 980

RESULT 3
BG049329 694 bp mRNA linear EST 25-JAN-2001
LOCUS OV1_18_E11_g1_A002 Ovary 1 (OV1) Sorghum bicolor cDNA, mRNA
DEFINITION sequence.
ACCESSION BG049329
VERSION BG049329.1 GI:12500932
KEYWORDS EST.
SOURCE Sorghum.
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 694)
Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt
,L.H.
An EST database from Sorghum: ovaries of varying immature stages
Unpublished (2000)
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860

Fax: 706 583 0210
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions
below Phred quality 16. The threshold for highest quality sequence
is 20.
Seq primer: PolyTWix
High quality sequence start: 36
High quality sequence stop: 689
POLYA-No.
Location/Qualifiers
1..694
/organism="Sorghum bicolor"
/db_xref="taxon:4558"
/clone_lib="Ovary 1 (OV1)"
/note="Organ: Mix of ovaries of varying immature stages
from 8-week-old plants; Vector: pBluescript II from lambda
zap II; Site_1: XhoI; Site_2: EcoRI; The library was made
from poly-A RNA in the cloning vector lambda ZAP II.
Clones to be sequenced were prepared by mass excision."
BASE COUNT 169 a 145 c 181 g 199 t
ORIGIN

Query Match 49.6%; Score 538.2; DB 12; Length 694;
Best Local Similarity 96.8%; Pred. No. 1.7e-155;
Matches 549; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
QY 520 GCTGGGATTTACTTACTGAACCCATCTGCTTGCACCGCATTCAGCTGAGGCCAACATCA 579
|||||
Db 1 GCTGGGATTTACTTACTGAACCCATCTGCTTGCACCGCATTCAGCTGAGGCCAACATCA 60
QY 580 ATTGAGAAAGAGGTCTTCCCTCAAAATTGACGCTGATCAACAGCTCTATGCAATGGTCTT 639
|||||
Db 61 ATTGAGAAAGAGGTCTTCCCTCAAAATTGACGCTGATCAGAGCTCTATGCAATGGTCTT 120
QY 640 CCAGGTTTTTGGATGGATGTTGGTCAGCCTAGGAGTACATTTACTGGCTTCGCTCTTAT 699
|||||
Db 121 CCAGGTTTTTGGATGGATGTTGGTCAGCAGCAAGGAGTACATTTACTGGCTTCGCTCTTAT 180
QY 700 CTAGACTCATTTAGGAAGAAATCAGCTGCCAGCTAGCTACTGAGGACACATGTTGTGGC 759
|||||
Db 181 CTAGACTCATTTAGGAAGAAATCAGCTGCCAGCTAGCTACTGAGGACACATGTTGTGGC 240
QY 760 AATGTCTGCTGTCATGAGAGCGCAAGATTGGAGAGGTTGCTGATTGGTCTCTGATGTC 819
|||||
Db 241 AATGTCTGCTGTCATGAGAGCGCAAGATTGGAGAGGTTGCTGATTGGTCTCTGATGTC 300
QY 820 GCCATTGGACCTGGGTGTTGTTGGAGAGCGCGGTGAGGCTTTCCCGCTCACTGTCATG 879
|||||
Db 301 GCCATTGGACCTGGGTGTTGTTGGAGAGGATGGCGTTCGCGCTCACTGTCATG 360
QY 880 CGCGGGTCCGTATCAAGAAGCATGCTTGCATCTCAACAGCATTTATCGGCTGGCACTCA 939
|||||
Db 361 CGTGTGTGCGCATCAAGAAGCATGCTTGCATCTCAACAGCATTTATCGGCTGGCACTCA 420
QY 940 ACTGTGTGTCATGGCACGATAGAGATATGACTATCTCTGGGAGGATGTTTCATGTC 999
|||||
Db 421 ACTGTGTGAAAATGGGACGGATAGAGATATGACTATCTCTGGGAGGATGTTTCATGTC 480
QY 1000 TGTGATGAGGTGTACAGCAATGGCGTGTGTTCTCCACATAAAGAGATCAAGTCAAGC 1059
|||||
Db 481 TGTGATGAGGTATACAGCAATGGCGTGTGTTCTCCACATAAAGAGATCAAGTCAAGC 540
QY 1060 ATTCTGAAGCCCTGAGATCGTCATGTGA 1086
|||||
Db 541 ATTCTGAAGCCCTGAGATCGTCATGTGA 567

RESULT 4
BE455732 836 bp mRNA linear EST 22-OCT-2001
LOCUS HVSMEG0019G12f Hordeum vulgare pre-anthesis spike EST library
DEFINITION HVCNDA0008 (white to yellow anther) Hordeum vulgare cDNA clone
HVSMEG0019G12f, mRNA sequence.

Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector, and regions below Phred quality 16. The threshold for highest quality sequence is 20. Three-prime sequences, which are obtained with PolyTmix or T7 sequencing primer, are presented as the reverse complement.
Seq primer: JEN REV
High quality sequence stop: 461
POLYA=No.

FEATURES

```

FEATURES
source
Location/Qualifiers
i. .524
/organism="Sorghum bicolor"
/cultivar="BFX623"
/db_xref="taxon:4558"
/clone_lib="Pathogen-infected compatible 1 (PCL1)"
/tissue_type="Leaves"
/dev_stage="4-week-old seedlings infected with
Colletotrichum graminicola"
/note="Vector: pBluescript II SK(-) from Lambda Zap II;
Site.1: XhoI; Site.2: EcoRI; Four-week-old sorghum
seedlings were sprayed with spore suspension prepared from
3-week-old FRM421, a sorghum isolate of the anthernose
pathogen Colletotrichum graminicola. Inoculated plants
were kept in a 25 C dark growth chamber with 100% relative
humidity for 24 hr, followed by 12/12 hr of light/dark
cycle at 25 C with 90% relative humidity for another 24
hr. All leaves were harvested and quick frozen with liquid
nitrogen and stored in a -80 C freezer. The library was
made from poly-A RNA in the cloning vector lambda Zap II.
Clones to be sequenced were prepared by mass excision.
WARNING: While most or all ESTs are expected to derive
from the host plant, no effort was made to eliminate ESTs
deriving from the pathogen."
138 a 1104 c 140 q 142 t
BASE COUNT

```

	Query Match	44.98;	Score 486;	DB 13;	Length 524;
	Best Local Similarity	96.1%;	Pred. No. 2.2e-139;		
	Matches 498;	Conservative 0;	Mismatches 20;	Indels 0;	Gaps 0;
QY	259	GTGGCCCTCTTGCTCTAGCAAGGACAAGCTTGCAGATGGATCTGGCAGCCATCTTT	318		
Db					
Db	7	GGTGGTCTCTTGCTCTAGCAAGGACAAGCTTGCAGATGGCTGGCGAGCCATCTTT	66		
QY	319	GTCTCAACAGTGGTATCATAGCGAATACCCATTCTGCTGAACATCATCAATTTCAAG	378		
Db	67	GTCTCAACAGTGGTATCATAGCGAATACCCATTCTGCTGAACATCATCAATTTCAAAA	126		
QY	379	TGTCATGGTGGTGAGGCAACAATATGGTCACTAAGGTGGATGAACCATCAAAATACGGT	438		
Db	127	TGTCATGGTGGTGAGGCAACAATATGGTCACTAAGGTGGATGAACCATCAAAATATGGT	186		
QY	439	GTTGTGGTTATGGAGGAGGCAACTGGCAGGTGGAAAGTTTGTGTGAGAGCCAAAAATA	498		
Db	187	GTTGTGGTTATGGAGGAGGCAACTGGCAGGTGGAAAGATTGTGTGAGAGCCAAAAGTA	246		
QY	499	TTTGTGGGTAAAGATCAATGTGGGATTACTTACTGAACCCATCTGCTCTTGACCGC	558		
Db	247	TTTGTGGGCACAGATCAATGTGGGATTACTTACTGAACCCATCTGCTCTTGACCGC	306		
QY	559	ATTGAGCTGAGGCCAACATCAATTGAGAAAGAGTCTTCCCTCAAAATGCACTGATCAA	618		
Db	307	ATTGAACTGAGGCCAACTCAATTGAGAAAGAGTCTTCCCTCAAAATGCACTGATCAG	366		
QY	619	CAGCTCTATGCAATGGTCTTCAGGTTTTTGGATGGATGTTGGTTCAGCCTAGGAGCTAC	678		
Db	367	AGGCTCTATGCAATGGTCTTCAGGTTTTTGGATGGATGTTGGTTCAGCCTAGGAGCTAC	426		
QY	679	ATTACTGGCTGGCTCTTTATCTAGACTCGATTAGGAAGAAATCAGCTGCCAAGCTAGCT	738		
Db	427	ATTACTGGCTGGCTCTTTATCTAGACTCTATTAGGAGAGATCAGCTGCCAAGCTAGCT	486		
QY	739	ACTGGAGCACATGTTGTTGCCAATGTGCTGGTGCATGA	776		

Db	487	ACTGGAACACATGTTGTGGCAATGCTGCTGTCATGA	524
RESULT	6		
LOCUS	BE405374		
DEFINITION	BE405374 597 bp mRNA linear EST 21-JUL-2000 WHE1213_F04_K072S Wheat etiolated seedling root cDNA library Triticum aestivum cDNA clone WHE1213_F04_K07, mRNA sequence.		
ACCESSION	BE405374		
VERSION	BE405374.1	GI:9364842	
KEYWORDS	EST.		
SOURCE	bread wheat.		
ORGANISM	Triticum aestivum		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poolideae; Triticeae; Triticum.		
AUTHORS	Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Han P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J., Seaton,C.L. and Tong,J.C.		
TITLE	The structure and function of the expressed portion of the wheat genomes		
JOURNAL	Unpublished (2000)		
COMMENT	Contact: Olin Anderson US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center 800 Buchanan Street, Albany, CA 94710, USA Tel: 5105959773 Fax: 5105959818 Email: oanderson@pw.usda.gov Sequence have been trimmed to remove vector sequence and low quality sequence with phred score less than 20 Seq primer: Strategene SK primer.		

FEATURES
SOURCE

```

1. .597
/organism="Triticum aestivum"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clones="WHE1213_F04_K07"
/clone_lib="wheat etiolated seedling root cDNA library"
/tissue_type="Root"
/dev_stage="Five day old etiolated seedling"
/lab_host="E. coli SOLR"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site_1: EcoRI; Site_2: XhoI; Seeds were surface-sterilized
, germinated and grown aseptically in the dark at room
temperature on filter paper with water, nystatin and
cefotaxime in covered crystallization dishes. Roots were
harvested. The tissue, total RNA, and poly(A) RNA were
prepared, a cDNA library was made, and the cDNA clones
were in vivo excised to give Bluescript phagemids in the
TJ Close lab (Choi, Close, Fenton) at the University of
California, Riverside. Plasmid DNA preparations and DNA
sequencing were performed in the OD Anderson lab (all
other authors)."

```

BASE COUNT	161 a	138 c	144 g	154 t
ORIGIN	Other authors).			

Query Match	44.5%	Score 483.4;	DB 10;	Length 597;
Best Local Similarity	88.1%	Pred. No. 1.6e-138;		
Matches 526;	Conservative	0;	Mismatches 71;	Indels 0;
				Gaps 0;

Qy	37	CGCCTTCGGCCCTTTGAC	CTCTGAGCTTCCCGAAACCCCTCGTGGATTTTGC	CAACAAGCC	96
Db	1	CGCCTTCGGCCCTCTC	ACGCTCAGCTTCCCCAAGCCCTCGTCGATTTTCG	CAACAAGCC	60
Qy	97	ATGATTCGACACAGATCG	CAAGCTTTGAAGAAGTTGGGGTCACAGAGTGGTTTGGCT	156	
Db	61	ATGATTCGACACAGATTG	AAGCCTTGAACAGCTTGGGGTTACAGAGTTATTTAGCC	120	
Qy	157	ATCAACTATCCCCAGAGGT	AATGATTAATTTCTTGAAGGACTTTGAGGATAAGCTTGGC	216	
Db	121	ATCAACTACCCCCAGAGGT	CATGATTAATTTCTTGAAGGATTTTGAAGATAAGCTTGGC	180	

QY	217	ATCAAAATTACATGCTCCCAAGAGACTGAGCCCTTAGGAACCGCTGGCCCTCTTGTGCTTA	276
Db	181	ATCAAAATCATGCTCCCAAGAGACTGAACCCCTGGGAAGTCTGGTCTCTTGGCCGTG	240
QY	277	GCAAGGCAAGGTTGGCGGATGCGATCGGCGAGCCATTCCTTTCCTCCTCAACAGTGAATGTC	336
DD	241	GCAAGGCAAGGTTGGCGGATGCGATCGGCGAGCCATTCCTTTCCTCCTCAACAGTGAATGTC	300
QY	337	ATAAGCAATACCCATTTGGTGAATCATCAAAATTTCAAAAGTCTCATGTGCTGAGGCA	396
Db	301	ATAAGCAATACCCATTTGGTGAATCATCAAAATTTCAAAAGTCTCATGTGCTGAGGCA	360
QY	397	ACAATATGCTCACTAAGTGGATGAACCATCAAAATACGGTGTGTGGTGTATGGAGAG	456
Db	361	ACAATATGCTCACTAAGTGGATGAACCATCAAAATACGGTGTGTGGTGTATGGAGAG	420
QY	457	GCAACTGGCAGGGTGGAAAGCTTTGTTGAGAGGCCAAAATATTTGTGGGTAAACAAGATC	516
Db	421	GCAACTGGCAGGGTGGAAAGCTTTGTTGAGAGGCCAAAATATTTGTGGGTAAACAAGATC	480
QY	517	AATGCTGGGATTTACTATCAAGCAACCATATGCTGCTGATGACCGCAATTCAGTACGACCAACA	576
Db	481	AATGCTGGGATTTACTATCAAGCAACCATATGCTGCTGATGACCGCAATTCAGTACGACCAACA	540
QY	577	TCAATTGAGAAAGAGGCTTCCCTCAAAATTCAGCTGATCAAGAGCTCATGCAATG	633
Db	541	TCAATTGAGAAAGAGGCTTCCCTCAAAATTCAGCTGATCAAGAGCTCATGCAATG	597

RESULT 7

BJ320079/c

LOCUS

DEFINITION

BJ320079 y. Ogihara unpublished cdNA library, EST 09-APR-2002

accestion

VERSION

KEYWORDS

SOURCE

ORGANISM

EST.

bread wheat.

Triticum aestivum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae; Triticeae; Triticum.

1 (bases 1 to 686)

Ogihara Y. and Murai K.

Expressed genes in Triticum aestivum

Unpublished (2002)

Contact: Tadasu Shin-1

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshin@genes.nig.ac.jp.

Location/Qualifiers

source

1. .686

```

FEATURES
source
1. .686
Location/Qualifiers
/organism="Triticum aestivum"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="whylf13906"
/tissue_lib="Y. Ogihara unpublished cDNA library, wh_yf"
/tissue_type="spikelet at early flowering"
/dev_stage="Feekes, scale 6"
/notes="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site_1: EcoRI; Site_2: XhoI; plants were grown under
hydroponic conditions at UC Davis, salt stressed for 12
hours, and for 7 days, then dissected and frozen (Akhunov
in J Dvorak Lab). Total RNA was prepared from sheath
tissue, equal quantities of RNA were pooled from the two
samples, polyA was purified from the pooled RNA, a cDNA
library was made, and the cDNA clones were in vivo
excised to give phagescript phagemids in the T7 Close lab
at the University of California, Riverside (Akhunov, Chiu

```

BASE COUNT		174 a		193 c		147 g		172 t	
ORIGIN									
Query Match		44.1%; Score 478.6; DB 13; Length 586;							
Best local similarity		86.9%; Pred. No. 5.5e-137;							
Matches 526; Conservative		0; Mismatches 79; Indels 0; Gaps 0;							
QY	482	TTGAGAGCCAAAATATTCTGGTAAACAGATCAATCTGCGATTTTACTTACTGTAACC	541						
Db	686	TAGAGAGCCAAAATATTCTAGGCACAGATCAATCTGCGATTTTACTTACTGTAACC	627						
QY	542	CATCTGCTCTTGACCGCATTTGAGCTGAGGCCAACATCAATTTGAGAGAGGTCTCCCTC	601						
Db	626	CGTCTGCTCGACCGCATCGAGTTAAAGCCAACTTCAATCGAGAAAGAGGTCTTTCCGC	567						
QY	602	AAATTGACGTGATCAACAGCTCTATGCAATGGTCTCTCCAGGTTTGTGATGATGTTG	661						
Db	566	GAATTGCTGCTGATCAGAAGCTCTACGCCATGGTCTCTCCAGGTTTGTGATGATGTTG	507						
QY	662	GTCAGCCTAGAGGACTCATTTACTTGGCTTGGCTTTTATCTAGACTCCATTTAGGAAGAAAT	721						
Db	506	GCCAGCCAAAGGACTCATTTACTTGGCTTGGCTTTTATCTAGACTCCATTTAGGAAGAAAT	447						
QY	722	CAGCTGCCAAGCTAGCTACTGGAGCACATGTTGTTGGCAATGTGCTGGTGCATCAGAGCG	781						
Db	446	CAGCTGCCAAGCTGGCCGCCGAGCACATATGTTGGGGAATGCTGCTGCACACAGCG	387						
QY	782	CCAGATTGGAGAGGTTGCTGATTGGTCTGATGTCGCCATTTGGACCTTGGGTGTTGTTG	841						
Db	386	CCAAGATTGGGGAGGTTGCTGATTGGTCTGATGTCGCTATTTGGACCTTGGATGCTGCTG	327						
QY	842	TGGAGGACGGCGTACAGCTTTCCGCTGCACTGCTCATCGCGGCGGTCGCTATCAAGAAC	901						
Db	326	TGGAGGATGTTGTTGAGGCTCTCCGATGTCAGGTTGATGCTGTTGGTGGCATTAAGAAGC	267						
QY	902	ATGCTTCGATCTCAACAGCATTTACGGCTGGCACTCAACTGTTGGTCAATGGGCACCGA	961						
Db	266	ATGCTTCGATCTCAACAGCATTTACGGCTGGCACTCAACTGTTGGTCAATGGGCACCGA	207						
QY	962	TAGACAATATGACTATTCCTGGGGAGGATGTTGATGTTGTTGATGAGGTGTACAGCAATG	1021						
Db	206	TAGACAATATGAGATCCTGGGAGAAGATGATGATGTTGATGAGGTGTACAGCAATG	147						
QY	1022	CGGCTGTTGTTCTCCACATAAAGAGATCAAGTCAACCATTCGACGCTTCAGATCTCA	1081						
Db	146	CGGCTGTTGTTCTCCGACAAAGAGATCAAGTCAAGCATCTCTGAGGCTTGAGATGCTCA	87						
QY	1082	TGTGA 1086							
Db	86	TGTCA 82							
RESULT 8		802 bp mRNA linear EST 11-APR-2001							
BG582583		802 bp mRNA linear EST 11-APR-2001							
LOCUS		EST484328 GVN Medicago truncatula cDNA clone PGVN-7015 5' end, mRNA							
DEFINITION		sequence.							
ACCESSION		BG582583							
VERSION		BG582583.1 GI:13597647							
KEYWORDS		EST.							
SOURCE		Barrel medic.							
ORGANISM		Medicago truncatula							
REFERENCE		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;							
AUTHORS		C.D., Van Aken,S., Utterback,T., Cho,J. and Fraser,C.M.							
TITLE		ESTs from one month old nitrogen-fixing root nodules of Medicago							

REFERENCE 1 (bases 1 to 651)
Sato,K., Saisho,D. and Takeda,K.
TITLE Barley EST sequencing project in NIG and Okayama Univ
JOURNAL Unpublished (2002)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES
source
1..651
/organism="Hordeum vulgare subsp. vulgare"
/cultivar="Haruna Nijo"
/db_xref="taxon:112509"
/clone="baal6n22"
/clone_lib="K. Sato unpublished cDNA library, cv. Haruna Nijo adult, heading stage top three leaves"
/tissue_type="top three leaves"
/dev_stage="adult, heading stage"
BASE COUNT 170 a 182 c 145 g 154 t
ORIGIN

Query Match 41.2%; Score 447.4; DB 10; Length 651;
Best Local Similarity 86.0%; Pred. No. 2.6e-127;
Matches 496; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 510 CAAGATCAATGCTGGGATTTACITACATGACCCATCTGCTTCACCCGATGAGCTGAG 569
|||||
DB 651 CAAGATCAATGACAGAAATTTACTTGTAAATCCATCTGCTTCACCCGATGAGTGA 592
|||||
QY 570 GCCAATCAATTCAGAAAGAGTCTCCCTCAATTCAGCTGATCAACAGCTCTATGC 629
|||||
DB 591 GCCAATCAATTCAGAAAGAGTCTCCCTCGAATTCAGCTGATCAACAGCTCTATGC 532
|||||
QY 630 AATGGTCTTCCAGGTTTTGGATGGATTTGGTGCAGCTAGGACTACATTAATCTGGCTT 689
|||||
DB 531 CATGGTCTTCCAGGCTTTGGATGGATATTGGCCAGCAAGGATTTACATCACTGGCTT 472
|||||
QY 690 CGGTCTTTATCTAGACTCGATTAGGAGAATATGAGTGCCTAGGCTGATCTATCTGGCTG 749
|||||
DB 471 CGGTCTTTATCTAGACTCGATTAGGAGAATATGAGTGCCTAGGCTGATCTATCTGGCTG 412
|||||
QY 750 TGTGTTGGCAATGCTGCTGATGAGCGCCCAAGATTGGAGAGTGTCTGATTTGG 809
|||||
DB 411 TTTGTTGGCAATGCTGCTGATGAGCGCCCAAGATTGGAGAGTGTCTGATTTGG 352
|||||
QY 810 TCCTGATGTCGCCATTTGACCTGGGTGTTGTGGAGCAGCGCTGAGGCTTTCCCGGTG 869
|||||
DB 351 CCCTGATGTTGCTATTGACCTGGATGCTGTTGGAGAGTGGTGTGAGGCTCTCCGATG 292
|||||
QY 870 CACTGTCATCGCGGCTGGCTATCAAGAAGCATGCTTGCATCTCAACAGCATTTATCGG 929
|||||
DB 291 CACGGTCATCGTGGTGTACGCATTAAAGAGCATGCGTGCATCTCAACAGCATCATCGG 232
|||||
QY 930 CTGGCACTCAACTGTTGGTCAATGGGACCGATAGAGATATGACTATCTCGGGGAGGA 989
|||||
DB 231 CTGGCACTCAACTGTTGGTCAATGGGACCGATAGAGATATGACTATCTCGGGGAGGA 172
|||||
QY 990 TGTTCATGTTGTGATGAGTGTACAGCAATGGCGGTGTTCTTCCACATTAAGAGAT 1049
|||||
DB 171 TGTACATGTTGGTACAGAGTGTACAGCAAGCGGGGTGTTCTTCCGCGCAAGGAGAT 112
|||||
QY 1050 CAAGTCAAGCATTTCTGAAGCGCTGAGATCGTCATGTGA 1086
|||||
DB 111 CAAGTCAAGCATTTCTGAAGCGCTGAGATCGTCATGTGA 75
|||||

RESULT 12
BE844694
LOCUS AD01B017 AD A. thaliana (Col-0 gll) linear EST 17-MAY-2001
DEFINITION AD01B017 AD A. thaliana (Col-0 gll) library enriched for

salt-induced transcripts; 10-14 day seedlings; 4h 160mM NaCl stress
Arabidopsis thaliana cDNA clone AD01B01 similar to T01007
mannose-1-phosphate guanylyltransferase (EC 2.7.7.13), mRNA
sequence.
BE844694
BE844694.1 GI:10277072
EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 794)
Gong,Z., Koiwa,H., Cushman,M.A., Ray,A., Bufford,D., Kore-Eda,S.,
Matsumoto,T.K., Zhu,J., Cushman,J.C., Bressan,R.A. and Hasegawa
,P.M.
Genes that are uniquely stress regulated in salt overly sensitive
(sos) mutants
Plant Physiol. 126 (1), 363-375 (2001)
21249177
COMMENT Contact: Cushman JC
Department of Biochemistry
University of Nevada
MS200, Reno, NV 89557-0014, USA
Tel: 775-784-1918
Fax: 775-784-1650
Email: jcushman@unr.edu
Plate: AD01 row: B column: 1
Seq primer: T7
POLYA=No.

FEATURES
source
1..794
/organism="Arabidopsis thaliana"
/strain="ecotype Columbia gll"
/db_xref="taxon:3702"
/clone="AD01B01"
/clone_lib="AD A. thaliana (Col-0 gll) library enriched
for salt-induced transcripts; 10-14 day seedlings; 4h
160mM NaCl stress"
/tissue_type="mixed leaf and root"
/dev_stage="10-14 day old"
/lab_host="Escherichia coli, TOP10F"
/note="Organ: whole seedlings; Vector: pT-Adv (CLONTECH);
Site.1: EcoRI; Site.2: EcoRI; AD Arabidopsis thaliana
library enriched for salt-induced transcripts from
wildtype (Col-0 gll) 10-14 day-old seedlings treated with
160 mM NaCl for 4 hours."
BASE COUNT 224 a 158 c 204 g 208 t
ORIGIN

Query Match 40.8%; Score 442.8; DB 12; Length 794;
Best Local Similarity 74.3%; Pred. No. 8.2e-126;
Matches 572; Conservative 0; Mismatches 197; Indels 1; Gaps 1;

QY 49 TTGACTCTCAGCTCCCGAAACCCCTCGGATTTTGCACCAAGCCCATGATCTGCAC 108
|||||
DB 20 TTGACTCTCAGTTCCCAAAAGCCCTGTGTGATTTTGTCTAAATAAACCCCATGATCTCAT 79
|||||
QY 109 CAGATCGAAGCTTTGAAAGAAAGTTGGGGTCACAGAGGTGGTTTGGCTATCAACTATCGC 168
|||||
DB 80 CAGATAGAGCTCTTAAGGCAGTTGGAGTTGATGAAGTGGTTTGGCCATCATTTATCAG 139
|||||
QY 169 CCAGAGGTAAATGATTAATTTCTTGAAGGACTTTTGAAGATAAGCTTGGCATCACAATTACA 228
|||||
DB 140 CCAGAGGTGATCTGAACCTCTTGAAGGACTTTGAGACCAAGCTGGAAATCAAAATCACT 199
|||||
QY 229 TGCTCCCAAGAGCTGAGCCCTTAGGAACCGCTGGCCCTCTCTCTCTAGCAAGGACAAG 288
|||||
DB 200 TGCTCACAAGAGACCGAGCCACTAGGTACCGCTGGTCTCTCTCTAGCGAGAGACAAA 259
|||||
QY 289 CTTGCGGATGGATCGCCAGAGCATTTCTTGTCTCAACAGTGTCTCATCAAGCGAATAC 348
|||||
DB 260 TTGCTTGTGATGTGGAGAGCCCTTCTTTGTTCTTAAACAGTGTGTGTAGTAGTAC 319
|||||

```

QY 349 CGATTGTGTAAGTCAATCAATTTACAGAGTGTATGTTGGTGGAGGCAAAATTTAGTGC 408
DB 320 CCTCTTAAAGAAATGCTTACAGTTTACAAAATCTCAGCGTGGAGAGCCCTCCATATGGTA 379
QY 409 ACTAGGTGGATGACCAATCAAAATACAGTGTGTGGTTATGAGGAGGCACTGGCAGG 468
DB 380 ACNAAAGGTGGATGACCGTGGAAATATGGAGTGGTTGTTATGGAAGAAAGCACTGGAAGA 439
QY 469 GTGGAAGGTTGTGACAGAGCCAAAATATTTCTGGGTAAACAAGATCAATGTCGCAAT 528
DB 440 GTGGAGAGTTGTGGAAGAGCCAAACTGTATGTAGGTAACAGATCAACGCTGGGAT 499
QY 529 TACTTACTGAACCCATCTGCTGACCGCATTTAGCTGAGCGCCAAACATCAATTTAGAGAA 588
DB 500 TATCTTCTGAACCCATCTGCTTCTGATAAGATGAGCAAGACCGATTTCAATCAAAA 559
QY 589 GAGGTCTTCCCTCAATTTGACAGCTGATCAACAGCTCTATGCAATGGTCTTCCAGGTTTT 648
DB 560 GAGACTTTTCTTAAGATTCGACAGCGGAGGCGCTCTATGCTATGCTGTACAGGTTT 619
QY 649 TGGATGGATTTGTCGACGCTAGGACATACATCTGCTTCCCTCTTATCTACACTCC 708
DB 620 TGGATGGATTTGTCGACGCTAGGACATACATCTGCTTCCCTCTTATCTACACTCC 679
QY 709 ATTAGGAAGAAATCACTGCGCAAGCTAGCTACTGAGGACATGTTGTTGGCAATGTGCTG 768
DB 680 CTTAGGAAGAAATCTCTGCGCAA-ATAACCACTGGGCGCACACATAGTTGGGAATTCGTT 738
QY 769 GTGCATGAGAGCGCAAGATTTGGAGAAGTTGTCTGATTTGGTCTGATGT 818
DB 739 TGTGACGAAGCGCTACAACTGGGGAAGGATGTTTGTATGGGACCAAGCCT 788

RESULT 13
BQ806351
LOCUS
DEFINITION
WHE3577_G06_N112s wheat developing grains cDNA library Triticum
aestivum CDNA CLONE WHE3577_G06_N11, mRNA sequence.
ACCESSION
BQ806351
VERSION
BQ806351.1 GI:22030560
KEYWORDS
EST.
SOURCE
bread wheat.
ORGANISM
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
; Triticeae; Triticum.
1 (bases 1 to 619)
Altenbach,S., Anderson,O.D., Chao,S., Chin,A., Close,T.J., Cronin
,K., Crossman,C., Penton,R.D., Lazo,G.R., Pham,J., Rausch,C.J.,
Wilson,C. and Woo,J.
The structure and function of the expressed portion of the wheat
genomes - Developing grains cDNA library
Unpublished (2002)
Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: Oanderson@pw.usda.gov
Sequences have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: SK primer.
Location/Qualifiers
1. .619
/organism="Triticum aestivum"
/cultivar="Butte 86"
/db_xref="taxon:4565"
/clone="WHE3577_G06_N11"
/clone_lib="Wheat developing grains cDNA library"
/tissue_type="whole grains"
/dev_stage="3-44 days post anthesis seed"
FEATURES
source

```

/lab_host="E. coli SOLR"
 /note="Vector: Lambda ZAP II, excised phagemid; Site 1:
 EcoRI; Plants were grown under six following different
 environmental regimes in greenhouse, Environment 1)
 24°C/17°C day/night, well-watered, with post-anthesis
 fertilizer, Environment 2) 24°C/17°C day/night,
 well-watered, without post-anthesis fertilizer,
 Environment 3) 37°C/17°C day/night, well-watered, with
 post-anthesis fertilizer, Environment 4) 37°C/17°C
 day/night, well-watered, without post-anthesis fertilizer,
 Environment 5) 37°C/17°C day/night plus drought, with
 post-anthesis fertilizer, Environment 6) 37°C/17°C
 day/night plus drought, without post-anthesis fertilizer,
 and frozen in liquid nitrogen, Environment 1 at 3, 5, 7,
 8, 10, 12, 16, 20, 24, 28, 32, 36, 40, 44 DPA Environment
 2 at 3, 5, 7, 8, 10, 12, 16, 20, 24, 28, 32, 36, 40, 44,
 DPA Environment 3 at 3, 5, 7, 8, 10, 12, 16, 20, 24, 28,
 32, 34 DPA Environment 4 at 3, 5, 7, 8, 10, 12, 16, 20, 24,
 28, 32, 34 DPA Environment 5 at 3, 5, 7, 8, 10, 12, 16,
 20, 24, 28, 30 DPA Environment 6 at 3, 5, 7, 8, 10, 12, 16,
 20, 24, 28, 30 DPA and total RNA was prepared by S.
 Altenbach and K. Cronin at USDA-ARS, Albany, CA. A cDNA
 library was made using poly (A) RNA, and the cDNA clones
 were in vivo excised to give pBluescript SK(-) phagemids
 in the T7 Close lab (Chin, Close, Penton) at the
 University of California, Riverside. Plasmid DNA
 preparations and DNA sequencing were performed in the OD
 Anderson lab (others)."

BASE COUNT 147 a 140 c 181 g 150 t 1 others
 ORIGIN

Query Match 40.3%; Score 438; DB 14; Length 619;

Best Local Similarity 85.3%; Pred. NO. 2.1e-124;

Matches 511; Conservative 0; Mismatches 86; Indels 2; Gaps 2;

QY 489 GCCAAAATATTTGGGTAAACAAGATCAATGCTGGGATTACTTACTGAACCATCTGT 548

DB 3 GGCACGAGGATTTGTAGGCAACAAGATCAATGCTGGAATTTACTCTTAAATCCCTCTGT 62

QY 549 CQPTGACCGCATTTGAGCTGAGCCACATCAATTTGAGAAGAGTCTTCCCTCAAATTCG 608

DB 63 CCTGGACCGCATCGAGTTAAAGCCAACTTCAATCGAGAAAGAGTCTTTCGCGAATTCG 122

QY 609 AGCTGATCAACACTCTATGCAATGCTCCCTCCAGGTTTGGATGATGTTGCTCAGCC 668

DB 123 TGTGATCAAGACTCTACGCCATGTTCTCCAGGTTTGGATGATTTGGCCAGCC 182

QY 669 TAGGGACTACATTACTGGCTTGGCTTTTATCTAGACTCGATTAGGAAGAAATCAGCTGC 728

DB 183 AAGGACTACATTACTGCGCTTGGCTTTTATCTAGACTCGATTAGGAAGAAATCAGCTGC 242

QY 729 CAAGCTTAGCTACTGGAGCACATCTTGTGCAATGCTGTCATGATGAGAGCGCCAGAT 788

DB 243 CAAGCTGGCCGCGGAGCACATATTGCGGAATGTCCTGTCGACGACGCGCC-AGAT 301

QY 789 TGCAGAGGTTCTGCTGATGCTGATGTCGCAATGTCGCAATGTCGCAATGTCGCAAT 848

DB 302 TGGGAGGGTTGCTTGAATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 361

QY 849 CGCGCTGAGGCTTTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 908

DB 362 TGTGTGAGGCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 421

QY 909 CATCTCAAAACAGCATTTATCGGCTGGCACTCAACTGTTGGTCAATGGCCAGGATAGAGAA 968

DB 422 CATCTCAAAACAGCATTTATCGGCTGGCACTCAACTGTTGGTCAATGGCCAGGATAGAGAA 481

QY 969 TATGACTATCTCGGGGAGGATGTTTCATGT-GTGTGATGAGGTGTACAGCAATGGCGGTG 1027

DB 482 TATGAGGATCTCGGGGAGGATGTATGATGTAGNGTACGAGGTCTACACAATGGCGGTG 541

QY 1026 TTGTTTCCCAACATAAAGACATCAAGTCAAGCATTTCTGCAAGCCTGACATCTCTCATCTCA 1086

vector sequence and contains a minimum of 100 bases of
phred value 20 or above. For more details on library
preparation and sequence analysis see
<http://www.genome.clemson.edu/projects/barley>. To order
this clone see <http://www.genome.clemson.edu/orders> Also
see Close TJ, Wing R, Kleinjofhs A, Wise R (2001)
Genetically and physically anchored EST resources for
barley genomics. Barley Genetics Newsletter 31:29-30.
(<http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html>)"

BASE COUNT	171 a	227 c	255 g	155 t	
ORIGIN					
Query Match	40.1%	Score 435;	DB 12;	Length 808;	
Best Local Similarity	75.5%	Pred. No. 2.2e-123;			
Matches 540;	Conservative 0;	Mismatches 175;	Indels 0;	Gaps 0;	
Qy	1	ATGAAGGCCCTCATCTGTGCGGGGTTTCGGAACCGCGCTTCGGCCTTTGACTCTGAGC	60		
Db					
Qy	59	ATGAAGGCCCTCATCTGTGCGGGGTTTCGGAACCGCGCTTCGGCCTTCGCTCAGC	118		
Db					
Qy	61	TTCCCGAAACCCCTCGTGGATTTCGAAACAGCCCATGATCTGCACAGATCGAAGCT	120		
Db					
Qy	119	GTGCCCAACCCCTCGTGGATTTCGCAACAGCCCATGATCTGCACAGATCGAAGCT	178		
Db					
Qy	121	TGGAAGAGTTGGGTACAGAGTGGTTTGGTATCAACTATGCCCCAGAGGTAATG	180		
Db					
Qy	179	CTGAAGATGTGGAGTCACAGAAGTTGCTCTGGCGATCAATTACGACGAGGTCATG	238		
Db					
Qy	181	ATTAATTTCTCAAGACATTTGAGATTAAGCTTGGCATCAACAAATACATCTCCCAAGAC	240		
Db					
Qy	239	CTCAACTTTCTCANGACTTTCGAGAGCAAGCTTGGCATCAAGATCACCTGTTCCTCCAGGAG	298		
Db					
Qy	241	ACTGAGCCTTAGGAACCGCTGGCCCTCTTGCTCTAGCAAGGCAAGCTTCGGGATGA	300		
Db					
Qy	299	ACAGACCAATGGAACCCCGGACCGCTGGCCCGCGGACAAAGCTCAGCAGCGA	358		
Db					
Qy	301	TCTGCCAGCCATTTCTTGTCTCAACAGTGTATCATAGGGAATACCCATTTGCTGAA	360		
Db					
Qy	359	TCCGCGAGCCTTCTTCTGCTCAACAGTGACGTGATCAGCGAGTACCCGTTCCGAGAG	418		
Db					
Qy	361	CTCATCAAAATTCACAAGTGTGCTGAGGCAACAAATATGGTCACTAAGGTGGAT	420		
Db					
Qy	419	CTCATCGAGTTCCACAGTCCCATGGCGGAGGCCACGATCATGCTACCAAGGTGGAC	478		
Db					
Qy	421	GAACCATCAAAATACGGTGTGCTTATGGAGAGGCAACTGGCAGGCTGGAAAGGTTT	480		
Db					
Qy	479	CAGCCTTCAGTACCGCTGCTGTGCTGACGAGGAGGACCCGGAAGGTGGGCGGTTT	538		
Db					
Qy	481	GTGAGAGCCAAAATATTTGTGGTAAACAGATCAATGCTGGGATTTACTTACTGAAC	540		
Db					
Qy	539	GTGGAGAAGCCCAAGGTGTCTGCGGCAACAAGATCAACCGCGGATCTACCTGCTGAC	598		
Db					
Qy	541	CAATCTGCTTCAGCCGATGAGCTGAGGCCAACAATTCAGAAAGAGGCTTTCCTT	600		
Db					
Qy	599	CCGTCGCTGAGCCGATGAGCTGAAGCGGACGTCCATCGAAGAGGAGGTGTCCCG	658		
Db					
Qy	601	CAATCTGAGCTGATCAACAGCTCTATGCAATGGTCTTCCAGCTTTTGGATGGATGTT	660		
Db					
Qy	659	CGTATCCCGGGGAAATAGGCTCTTCGCCATGCTGCTGCCGGTTCGGATGGACCAT	718		
Db					
Qy	661	GGTCAGCTAGGAGCTACATTACTGCTTGGCTTTATCTATAGACTCGATTAGGA	715		
Db					
Qy	719	CGGACCGCAAGGAGCTTACTTACCGGCTGAGGCTCTACCTTGAAGTCTGAGGA	773		
Db					

Search completed: November 27, 2002, 04:03:48
Job time : 1994 SECS

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 27, 2002, 04:05:27 ; Search time 67 Seconds
(without alignments)
717.962 Million cell updates/sec

Title: us-09-374-967-2

Perfect score: 1864

Sequence: 1 MKALILVGGFGRRLRLTL.....GVLPKHKSSILKPEIVM 361

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002.*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1835	98.4	361	21	AAAG34077
2	1681	90.2	361	21	AAAG20009
3	1681	90.2	361	21	AAAG46459
4	1681	90.2	361	21	AAV77936
5	1680	90.1	361	20	AAV08879
6	1671	89.6	329	21	AAAG34078
7	1540	82.6	302	21	AAAG34079
8	1517	81.4	329	21	AAAG20010
9	1517	81.4	329	21	AAAG46460
10	1402	75.2	302	21	AAAG20011

11	1402	75.2	302	21	AAAG46461	Arabidopsis thalia
12	1183.5	63.5	360	21	AAAB41757	Human OREF ORF1521
13	1183.5	63.5	360	21	AAAY79210	Human transferrase
14	1175	63.0	363	22	AAAM52357	NBP-hexose pyropho
15	1173	62.9	369	22	ABBS7869	Drosophila melanog
16	1173	62.9	369	22	ABB67221	Drosophila melanog
17	1160	62.2	387	22	AAAG95807	Human protein sequ
18	682	36.6	152	21	AAAG27153	Zea mays protein f
19	669	35.9	143	21	AAAG27154	Zea mays protein f
20	631.5	33.9	240	22	AAU23276	Novel human enzyme
21	592.5	31.8	420	22	AAAB92548	Human protein sequ
22	592.5	31.8	420	22	AAAB94085	Human protein sequ
23	554	29.7	411	21	AAAG17545	Arabidopsis thalia
24	554	29.7	415	21	AAAG17544	Arabidopsis thalia
25	550	29.5	411	21	AAAG38777	Arabidopsis thalia
26	550	29.5	415	21	AAAG38776	Arabidopsis thalia
27	545	29.2	402	21	AAAG17546	Arabidopsis thalia
28	541	29.0	402	21	AAAG38778	Arabidopsis thalia
29	529.5	28.4	415	21	AAAG33398	Zea mays protein f
30	529.5	28.4	431	21	AAAG33397	Zea mays protein f
31	521	28.0	448	22	ABBB62932	Drosophila melanog
32	520.5	27.9	403	21	AAAG33399	Zea mays protein f
33	512.5	27.5	111	21	AAAG22702	Zea mays protein f
34	493.5	26.5	385	21	AAAG29705	Arabidopsis thalia
35	488	26.2	359	22	AAAG81218	Mycobacterium tube
36	484.5	26.0	375	21	AAAG29706	Arabidopsis thalia
37	478	25.6	416	22	AAAB96271	Putative nucleosid
38	465	24.9	92	21	AAAG32844	Zea mays protein f
39	465	24.9	152	21	AAAG32843	Zea mays protein f
40	444	23.8	361	22	AAAB96151	Putative P. abyssi
41	442.5	23.7	362	22	AAAB79379	Corynebacterium gl
42	440	23.6	340	21	AAAG12777	Arabidopsis thalia
43	431.5	23.1	348	22	AAAG90571	C. glutamicum prote
44	411.5	22.1	346	21	AAAG29707	Arabidopsis thalia
45	409	21.9	306	21	AAAG12778	Arabidopsis thalia

ALIGNMENTS

RESULT 1

AAAG34077
ID AAAG34077 standard; Protein; 361 AA.
XX
AC AAAG34077;
XX
DT 18-OCT-2000 (first entry)
XX
DE Zea mays protein fragment SEQ ID NO: 41410.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.
XX
OS Zea mays subsp. mays.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 03-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130691.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 11-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136302.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137242.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147933.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.

PR	14-OCT-1999;	99US-0159331.	PN	EP1033405-A2.
PR	14-OCT-1999;	99US-0159637.	XX	
PR	14-OCT-1999;	99US-0159638.	PD	
PR	18-OCT-1999;	99US-0159584.	XX	06-SEP-2000.
PR	21-OCT-1999;	99US-0160741.	PF	
PR	21-OCT-1999;	99US-0160767.	XX	25-FEB-2000; 2000EP-0301439.
PR	21-OCT-1999;	99US-0160768.		
PR	21-OCT-1999;	99US-0160770.	PR	25-FEB-1999; 99US-0121825.
PR	21-OCT-1999;	99US-0160814.	PR	05-MAR-1999; 99US-0123180.
PR	21-OCT-1999;	99US-0160815.	PR	09-MAR-1999; 99US-0123548.
PR	22-OCT-1999;	99US-0160980.	PR	23-MAR-1999; 99US-0125788.
PR	22-OCT-1999;	99US-0160981.	PR	25-MAR-1999; 99US-0126264.
PR	22-OCT-1999;	99US-0160989.	PR	29-MAR-1999; 99US-0126785.
PR	22-OCT-1999;	99US-0160989.	PR	01-APR-1999; 99US-0127462.
PR	25-OCT-1999;	99US-0161404.	PR	06-APR-1999; 99US-0128234.
PR	25-OCT-1999;	99US-0161405.	PR	16-APR-1999; 99US-0128714.
PR	25-OCT-1999;	99US-0161406.	PR	16-APR-1999; 99US-0129845.
PR	26-OCT-1999;	99US-0161359.	PR	19-APR-1999; 99US-0130077.
PR	26-OCT-1999;	99US-0161360.	PR	21-APR-1999; 99US-0130449.
PR	26-OCT-1999;	99US-0161361.	PR	23-APR-1999; 99US-0130510.
PR	28-OCT-1999;	99US-0161920.	PR	23-APR-1999; 99US-0130891.
PR	28-OCT-1999;	99US-0161992.	PR	28-APR-1999; 99US-0131449.
PR	28-OCT-1999;	99US-0161993.	PR	30-APR-1999; 99US-0132048.
PR	29-OCT-1999;	99US-0162142.	PR	30-APR-1999; 99US-0132407.
Query Match 98.4%; Score 1835; DB 21; Length 361;				
Best Local Similarity 98.3%; Pred. No. 1.3e-193;				
Matches 355; Conservative 3; Mismatches 3; Indels 0; Gaps 0;				
QY	1	MKALILVGGFGTRLRPLTLSPFKPLVDFANKPMLHQIEALKEGVTEVVLAINRPEVM 60		
Db	1	MKALILVGGFGTRLRPLTLSPFKPLVDFANKPMLHQIEALKEGVTEVVLAINRPEVM 60		
QY	61	INFLKDFEDKLGITITCSQETPLGTAGPLALARDKLDGSGQPFVFLNSDVISEYPAE 120		
Db	61	INFLKDFEDKLGITITCSQETPLGTAGPLALARDKLDGSGEPFVFLNSDVISEYPAE 120		
QY	121	LIFPHKCHGGEATIMTKVDPSKYGVVVMEEATGRVERFEKPKIFVGNKINAGIYLLN 180		
Db	121	LIFPHKSHGGEATIMTKVDPSKYGVVVMEEATGRVERFEKPKIFVGNKINAGIYLLN 180		
QY	181	PSVLDRELPTSTSEKEVFFQIAADQOLYAMVLPGFWMVQQRDQITGLRLYLDLSIRKK 240		
Db	181	PSVLDRELPTSTSEKEVFFQIAADQOLYAMVLPGFWMVQQRDQITGLRLYLDLSIRKK 240		
QY	241	SAAKLATGARHVGNVNLVHESAKIGEGCLIGPDVAIGPGCVVVEDGVRLSRCVTVMRGVRIKK 300		
Db	241	SAARLAAGARHVGNVNLVHESAKIGEGCLIGPDVAIGPGCVVVDGVRLSRCVTVMRGVRIKK 300		
QY	301	HACISNSIIGHSTVGQWARIENMTILGEDVHVCDVYSNGGVVLPHPHKEIKSSILKPEIV 360		
Db	301	HACISNSIIGHSTVGQWARIENMTILGEDVHVCDVYSNGGVVLPHPHKEIKSSILKPEIV 360		
QY	361	M 361		
Db	361	M 361		
RESULT 2				
AAG20009				
ID	AAG20009 standard; Protein; 361 AA.			
XX				
AC	AAG20009;			
XX				
DT	17-OCT-2000 (first entry)			
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 22030.			
XX				
KW	Protein identification; signal transduction pathway; metabolic pathway;			
KW	hybridisation assay; genetic mapping; gene expression control; promoter;			
KW	termination sequence.			
XX	Arabidopsis thaliana.			
XX				

```
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144083.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144603.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0148171.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147933.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149175.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151348.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.

PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 07-OCT-1999; 99US-0157753.
PR 07-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 90.2%; Score 1681; DB 21; Length 361;
Best Local Similarity 88.6%; Pred. No. 1.4e-176;
Matches 320; Conservative 22; Mismatches 19; Indels 0; Gaps 0;

QY 1 MKALILVGEGTRLRPLTSLSPKPLVDFAFKPMILHQLHIEALKEVGYTEVYVLAINYRPEYM 60
DD 1 MKALILVGGFTRLRPLTSLSPKPLVDFAFKPMILHQLHIEALKEVGYTEVYVLAINYRPEYM 60
QY 61 INFLKDFEDKLGTTTCQSQETPLCTAGDIALARDKLIADGSGQDPFFVLNSDVISSEYFAE 120
DB 61 LNPLKDFETKLEIKITCSQETPLCTAGDIALARDKLIADGSGQDPFFVLNSDVISSEYFAE 120
QY 121 LKIFHKHCHGEATIMVTKVDEPSKYGVVWMEATGRVEREVEPKIFVGNKINAGIYLLN 180
DB 121 MLEFHSHGGEASIMVTKVDEPSKYGVVWMEATGRVEREVEPKIFVGNKINAGIYLLN 180
QY 181 PSVLDRIELRPTISIEKEVFPQIAADOOOLYAMVLPFGFMDVQOPRDYITGLRLYLDSTRKK 240
DB 181 PSVLDRIELRPTISIEKEVFPQIAADOOOLYAMVLPFGFMDVQOPRDYITGLRLYLDSTRKK 240
QY 241 SAAKLATGAHVGNVNLVHESAKITGECCLIGDPDVAIGPGCVVEDCVRLSRTVMRGVRIKK 300
DB 241 SPAKLTSGPHVGNVNLVDETTATIGEGCLIGDPDVAIGPGCVLIVSGVRLSRTVMRGVRIKK 300
QY 301 HACISNSIIGWSTVYGQWARIENMTILGDDVHYVCDEVYNSGVVLPKHEKTSSTLRPEIV 360
DB 301 HACISNSIIGWSTVYGQWARIENMTILGDDVHYVCDEVYNSGVVLPKHEKTSSTLRPEIV 360
QY 361 M 361
DB 361 M 361
```

RESULT 3
AAG46459
ID AAG46459 standard; Protein; 361 AA.
XX
AC AAG46459;
DT
DT 18-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 58452.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.

```
DR 20-AUG-1999; 99US-0149723.
FR 20-AUG-1999; 99US-0149929.
FR 23-AUG-1999; 99US-0149902.
FR 23-AUG-1999; 99US-0149930.
FR 25-AUG-1999; 99US-0150566.
FR 26-AUG-1999; 99US-0150884.
FR 27-AUG-1999; 99US-0151065.
FR 27-AUG-1999; 99US-0151066.
FR 27-AUG-1999; 99US-0151080.
FR 30-AUG-1999; 99US-0151303.
FR 31-AUG-1999; 99US-0151438.
FR 01-SEP-1999; 99US-0151930.
FR 07-SEP-1999; 99US-0152363.
FR 10-SEP-1999; 99US-0153070.
FR 13-SEP-1999; 99US-0153758.
FR 15-SEP-1999; 99US-0154018.
FR 16-SEP-1999; 99US-0154039.
FR 20-SEP-1999; 99US-0154779.
FR 22-SEP-1999; 99US-0155139.
FR 23-SEP-1999; 99US-0155486.
FR 24-SEP-1999; 99US-0155659.
FR 28-SEP-1999; 99US-0156458.
FR 20-SEP-1999; 99US-0156596.
FR 04-OCT-1999; 99US-0157117.
FR 05-OCT-1999; 99US-0157753.
FR 06-OCT-1999; 99US-0157865.
FR 07-OCT-1999; 99US-0158029.
FR 08-OCT-1999; 99US-0158212.
FR 12-OCT-1999; 99US-0158369.
FR 13-OCT-1999; 99US-0159293.
FR 13-OCT-1999; 99US-0159294.
FR 13-OCT-1999; 99US-0159295.
FR 14-OCT-1999; 99US-0159329.
FR 14-OCT-1999; 99US-0159330.
DR 14-OCT-1999; 99US-0159331.
FR 14-OCT-1999; 99US-0159637.
FR 14-OCT-1999; 99US-0159638.
DR 18-OCT-1999; 99US-0159584.
FR 21-OCT-1999; 99US-0160741.
FR 21-OCT-1999; 99US-0160767.
DR 21-OCT-1999; 99US-0160768.
FR 21-OCT-1999; 99US-0160770.
FR 21-OCT-1999; 99US-0160814.
FR 21-OCT-1999; 99US-0160815.
FR 22-OCT-1999; 99US-0160980.
FR 22-OCT-1999; 99US-0160981.
DR 22-OCT-1999; 99US-0160989.
FR 23-OCT-1999; 99US-0161404.
FR 25-OCT-1999; 99US-0161405.
FR 25-OCT-1999; 99US-0161406.
FR 26-OCT-1999; 99US-0161359.
FR 26-OCT-1999; 99US-0161360.
FR 26-OCT-1999; 99US-0161361.
FR 28-OCT-1999; 99US-0161920.
FR 28-OCT-1999; 99US-0161992.
FR 28-OCT-1999; 99US-0161993.
FR 29-OCT-1999; 99US-0162142.

Query Match 90.2%; Score 1681; DB 21; Length 361;
Best Local Similarity 88.6%; Pred. No. 1.4e-176;
Matches 320; Conservative 22; Mismatches 19; Indels 0; Gaps 0;

Qy 1 MKALILVGGCTGRLRDLTSLSPKPLVDFAKPMILHQTALKEVGVTEVLAINRPEVM 60
Dd 1 MKALILVGGCTGRLRDLTSLSPKPLVDFAKPMILHQTALKEVGVTEVLAINRPEVM 60

Qy 61 INFLKDFEDKLGITITCSQTERLGTAGLALARDKLDGSGQPFVFLNSDVISEYPAE 120
Dd 61 LNFLKDFETKLEIKTCSQTERLGTAGLALARDKLDGSGEFVFLNSDVISEYPLKE 120

Qy 121 LIKPHKHGGGATIMVTKVDPSYGVVVMEEATGRVERFVEKPKIFVGNKINAGIYLLN 180
Dd 121 MLEPHKHGGGASIMVTKVDPSYGVVVMEEATGRVEKPKIFVGNKINAGIYLLN 180
```

```
QY 181 PSVLDRIELRPTSIKEVEFFQIAADQOLXAMYLPGFWMVDYQGPDXITIGURLYDLSIRKK 240
Dd 181 PSVLDRIELRPTSIKEVEFFQIAADQOLXAMYLPGFWMVDYQGPDXITIGURLYDLSIRKK 240
QY 241 SAAKLATGAHVGNVNLVHESAKIGEGCLIGDPNVAIGPGCVVDEGVRLSRCTVMRGVRIKK 300
Dd 241 SPAKLITSGPHIVGNVLVDETATIGEGCLIGDPNVAIGPGCVVDEGVRLSRCTVMRGVRIKK 300
QY 301 HACISNSIICWSTVQWARIENKMTILGEDVHVCDVYSGVGVLPKHETKSSILKPEIV 360
Dd 301 HACISNSIICWSTVQWARIENKMTILGEDVHVCDVYSGVGVLPKHETKSSILKPEIV 360
QY 361 M 361
Dd 361 M 361

RESULT 4
AAV77936
ID AAV77936 standard; protein; 361 AA.
XX
AC AAV77936;
XX
DT 14-JUN-2000 (first entry)
XX
DE A. thaliana environmental stress tolerance related protein.
KW Environmental stress; plant; transgenic plant; anaerobic; flooding; cold;
KW dehydration; drought; heat stress; salinity; osmotolerance.
OS Arabidopsis thaliana.
XX
PN WO200008187-A2.
XX
PD 17-FEB-2000.
XX
PF 04-AUG-1999; 99WO-EP05652.
XX
PR 04-AUG-1998; 98EP-0202634.
XX
PA (VLAA-) YLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
PI Lee JH, Verbruggen N;
XX
XX
DR WPI: 2000-205726/18.
DR N-PEDB; AAZ98316.
XX
Isolation of polynucleic acids useful for producing transgenic plant by
isolating genes involved in tolerance to environmental stress -
Claim 12; Page 127-129; 312pp; English.
The invention relates to isolation of coding sequences and/or genes
involved in tolerance to environmental stress in plants. The sequences
(AAZ98305-298365) are useful for producing a transgenic plant having
enhanced tolerance or resistance to environmental stress conditions such
as anaerobic, flooding, cold, dehydration, drought, heat stress or
salinity. This is useful for producing improved yield, growth,
development and productivity under environmental stress conditions, and
also provides growth of crops in areas where they cannot grow without
the induced osmotolerance. Sequences AAV77925-984 represent polypeptide
sequences from A. thaliana that are encoded by the genes involved in
environmental stress tolerance.
Sequence 361 AA;
Query Match 90.2%; Score 1681; DB 21; Length 361;
Best Local Similarity 88.6%; Pred. No. 1.4e-176;
Matches 320; Conservative 22; Mismatches 19; Indels 0; Gaps 0;
Qy 1 MKALILVGGCTGRLRDLTSLSPKPLVDFAKPMILHQTALKEVGVTEVLAINRPEVM 60
Dd 1 MKALILVGGCTGRLRDLTSLSPKPLVDFAKPMILHQTALKEVGVTEVLAINRPEVM 60
```


PR	30-APR-1999;	99US-0132407.	PR	21-JUL-1999;	99US-0145088.
PR	04-MAY-1999;	99US-0132484.	PR	22-JUL-1999;	99US-0145085.
PR	05-MAY-1999;	99US-0132485.	PR	22-JUL-1999;	99US-0145087.
PR	06-MAY-1999;	99US-0132486.	PR	22-JUL-1999;	99US-0145089.
PR	07-MAY-1999;	99US-0132487.	PR	22-JUL-1999;	99US-0145192.
PR	07-MAY-1999;	99US-0132863.	PR	23-JUL-1999;	99US-0145145.
PR	11-MAY-1999;	99US-0134256.	PR	23-JUL-1999;	99US-0145218.
PR	14-MAY-1999;	99US-0134218.	PR	23-JUL-1999;	99US-0145224.
PR	14-MAY-1999;	99US-0134219.	PR	26-JUL-1999;	99US-0145276.
PR	14-MAY-1999;	99US-0134221.	PR	27-JUL-1000;	99US-0145913.
PR	14-MAY-1999;	99US-0134370.	PR	27-JUL-1999;	99US-0145918.
PR	18-MAY-1999;	99US-0134768.	PR	27-JUL-1999;	99US-0145919.
PR	19-MAY-1999;	99US-0134941.	PR	28-JUL-1999;	99US-0145951.
PR	20-MAY-1999;	99US-0135124.	PR	02-AUG-1999;	99US-0146386.
PR	21-MAY-1999;	99US-0135353.	PR	02-AUG-1999;	99US-0146388.
PR	24-MAY-1999;	99US-0135629.	PR	02-AUG-1999;	99US-0146389.
PR	25-MAY-1999;	99US-0136021.	PR	03-AUG-1999;	99US-0147038.
PR	27-MAY-1999;	99US-0136392.	PR	04-AUG-1999;	99US-0147204.
PR	28-MAY-1999;	99US-0136782.	PR	04-AUG-1999;	99US-0147302.
PR	01-JUN-1999;	99US-0137222.	PR	05-AUG-1999;	99US-0147192.
PR	03-JUN-1999;	99US-0137528.	PR	05-AUG-1999;	99US-0147260.
PR	04-JUN-1999;	99US-0137502.	PR	06-AUG-1999;	99US-0147303.
PR	07-JUN-1999;	99US-0137724.	PR	06-AUG-1999;	99US-0147416.
PR	08-JUN-1999;	99US-0138094.	PR	09-AUG-1999;	99US-0147493.
PR	10-JUN-1999;	99US-0138540.	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	99US-0138847.	PR	10-AUG-1999;	99US-0148171.
PR	14-JUN-1999;	99US-0139119.	PR	11-AUG-1999;	99US-0148319.
PR	16-JUN-1999;	99US-0139452.	PR	12-AUG-1999;	99US-0148341.
PR	16-JUN-1999;	99US-0139453.	PR	13-AUG-1999;	99US-0148565.
PR	17-JUN-1999;	99US-0139492.	PR	13-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	99US-0139454.	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139455.	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139456.	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-0139457.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139458.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139459.	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	99US-0139460.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139461.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139462.	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	99US-0139463.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139750.	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151066.
PR	21-JUN-1999;	99US-0139817.	PR	27-AUG-1999;	99US-0151080.
PR	22-JUN-1999;	99US-0139899.	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	99US-0140353.	PR	31-AUG-1999;	99US-0151438.
PR	23-JUN-1999;	99US-0140354.	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	99US-0140695.	PR	07-SEP-1999;	99US-0152363.
PR	28-JUN-1999;	99US-0140823.	PR	10-SEP-1999;	99US-0153070.
PR	29-JUN-1999;	99US-0140991.	PR	13-SEP-1999;	99US-0153758.
PR	30-JUN-1999;	99US-0141287.	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0141842.	PR	16-SEP-1999;	99US-0154039.
PR	01-JUL-1999;	99US-0142154.	PR	20-SEP-1999;	99US-0154779.
PR	02-JUL-1999;	99US-0142055.	PR	22-SEP-1999;	99US-0155139.
PR	06-JUL-1999;	99US-0142390.	PR	23-SEP-1999;	99US-0155486.
PR	08-JUL-1999;	99US-0142803.	PR	24-SEP-1999;	99US-0155659.
PR	09-JUL-1999;	99US-0142920.	PR	28-SEP-1000;	99US-0156458.
PR	12-JUL-1999;	99US-0142977.	PR	29-SEP-1999;	99US-0156596.
PR	13-JUL-1999;	99US-0143542.	PR	04-OCT-1999;	99US-0157117.
PR	14-JUL-1999;	99US-0143624.	PR	05-OCT-1999;	99US-0157753.
PR	15-JUL-1999;	99US-0144005.	PR	06-OCT-1999;	99US-0157865.
PR	16-JUL-1999;	99US-0144085.	PR	07-OCT-1999;	99US-0158029.
PR	16-JUL-1999;	99US-0144086.	PR	08-OCT-1999;	99US-0158232.
PR	19-JUL-1999;	99US-0144325.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144331.	PR	13-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-0144333.	PR	13-OCT-1999;	99US-0159295.
PR	19-JUL-1999;	99US-0144334.	PR	14-OCT-1999;	99US-0159329.
PR	20-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159330.
PR	20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159331.
PR	20-JUL-1999;	99US-0144632.	PR	14-OCT-1999;	99US-0159637.
PR	20-JUL-1999;	99US-0144884.	PR	14-OCT-1999;	99US-0159638.
PR	21-JUL-1999;	99US-0144814.	PR	18-OCT-1999;	99US-0159584.
PR	21-JUL-1999;	99US-0145086.	PR	21-OCT-1999;	99US-0160741.

PR 16-JUL-1999;	99US-0144085.	PR 07-OCT-1999;	99US-0158029.
PR 16-JUL-1999;	99US-0144086.	PR 08-OCT-1999;	99US-0158232.
PR 19-JUL-1999;	99US-0144325.	PR 12-OCT-1999;	99US-0158369.
PR 10-JUL-1999;	99US-0144331.	PR 13-OCT-1999;	99US-0159209.
PR 19-JUL-1999;	99US-0144332.	PR 13-OCT-1999;	99US-0159294.
PR 19-JUL-1999;	99US-0144333.	PR 13-OCT-1999;	99US-0159295.
PR 19-JUL-1999;	99US-0144334.	PR 14-OCT-1999;	99US-0159329.
PR 19-JUL-1999;	99US-0144335.	PR 14-OCT-1999;	99US-0159330.
PR 20-JUL-1999;	99US-0144352.	PR 14-OCT-1999;	99US-0159331.
PR 20-JUL-1999;	99US-0144632.	PR 14-OCT-1999;	99US-0159637.
PR 20-JUL-1999;	99US-0144884.	PR 14-OCT-1999;	99US-0159638.
PR 21-JUL-1999;	99US-0144814.	PR 18-OCT-1999;	99US-0159584.
PR 21-JUL-1999;	99US-0145086.	PR 21-OCT-1999;	99US-0160741.
PR 21-JUL-1999;	99US-0145088.	PR 21-OCT-1999;	99US-0160767.
PR 22-JUL-1999;	99US-0145085.	PR 21-OCT-1999;	99US-0160768.
PR 22-JUL-1999;	99US-0145087.	PR 21-OCT-1999;	99US-0160770.
PR 22-JUL-1999;	99US-0145089.	PR 21-OCT-1999;	99US-0160814.
PR 22-JUL-1999;	99US-0145192.	PR 21-OCT-1999;	99US-0160815.
PR 23-JUL-1999;	99US-0145145.	PR 22-OCT-1999;	99US-0160980.
PR 23-JUL-1999;	99US-0145218.	PR 22-OCT-1999;	99US-0160981.
PR 23-JUL-1999;	99US-0145224.	PR 22-OCT-1999;	99US-0160989.
PR 26-JUL-1999;	99US-0145276.	PR 25-OCT-1999;	99US-0161404.
PR 27-JUL-1999;	99US-0145513.	PR 25-OCT-1999;	99US-0161405.
PR 27-JUL-1999;	99US-0145918.	PR 25-OCT-1999;	99US-0161406.
PR 27-JUL-1999;	99US-0145919.	PR 26-OCT-1999;	99US-0161359.
PR 28-JUL-1999;	99US-0145931.	PR 26-OCT-1999;	99US-0161360.
PR 02-AUG-1999;	99US-0146386.	PR 26-OCT-1999;	99US-0161361.
PR 02-AUG-1999;	99US-0146388.	PR 28-OCT-1999;	99US-0161920.
PR 02-AUG-1999;	99US-0146389.	PR 28-OCT-1999;	99US-0161922.
PR 03-AUG-1999;	99US-0147038.	PR 28-OCT-1999;	99US-0161993.
PR 04-AUG-1999;	99US-0147204.	PR 29-OCT-1999;	99US-0162142.
PR 04-AUG-1999;	99US-0147302.		
PR 05-AUG-1999;	99US-0147192.		
PR 05-AUG-1999;	99US-0147260.		
PR 06-AUG-1999;	99US-0147303.		
PR 06-AUG-1999;	99US-0147416.		
PR 09-AUG-1999;	99US-0147493.		
PR 09-AUG-1999;	99US-0147935.		
PR 10-AUG-1999;	99US-0148171.		
PR 12-AUG-1999;	99US-0148319.		
PR 13-AUG-1999;	99US-0148341.		
PR 13-AUG-1999;	99US-0148565.		
PR 13-AUG-1999;	99US-0148684.		
PR 16-AUG-1999;	99US-0149368.		
PR 17-AUG-1999;	99US-0149175.		
PR 18-AUG-1999;	99US-0149426.		
PR 20-AUG-1999;	99US-0149722.		
PR 20-AUG-1999;	99US-0149723.		
PR 20-AUG-1999;	99US-0149929.		
PR 23-AUG-1999;	99US-0149902.		
PR 23-AUG-1999;	99US-0149930.		
PR 23-AUG-1999;	99US-0150566.		
PR 26-AUG-1999;	99US-0150884.		
PR 27-AUG-1999;	99US-0151065.		
PR 27-AUG-1999;	99US-0151066.		
PR 27-AUG-1999;	99US-0151080.		
PR 30-AUG-1999;	99US-0151303.		
PR 31-AUG-1999;	99US-0151438.		
PR 01-SEP-1999;	99US-0151930.		
PR 07-SEP-1999;	99US-0152363.		
PR 10-SEP-1999;	99US-0153070.		
PR 13-SEP-1999;	99US-0153758.		
PR 15-SEP-1999;	99US-0154018.		
PR 16-SEP-1999;	99US-0154039.		
PR 20-SEP-1999;	99US-0154779.		
PR 22-SEP-1999;	99US-0155139.		
PR 23-SEP-1999;	99US-0155139.		
PR 24-SEP-1999;	99US-0155486.		
PR 24-SEP-1999;	99US-0155659.		
PR 28-SEP-1999;	99US-0156458.		
PR 29-SEP-1999;	99US-0156596.		
PR 04-OCT-1999;	99US-0157117.		
PR 05-OCT-1999;	99US-0157753.		
PR 06-OCT-1999;	99US-0157865.		

Query Match 82.6%; Score 1540; DB 21; Length 302;
Best Local Similarity 98.0%; Pred. No. 3.9e-161;
Matches 296; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 60 MINFLKDFEDKLGITITCSQETPLCTAGLALARDKLDGSGQPFVFLNSDVISEYPPFA 119
|||||

Db 1 MINFLKDFEDKLGITITCSQETPLCTAGLALARDKLDGSGQPFVFLNSDVISEYPPFA 60
|||||

Qy 120 ELIKFKHCHGGEATIMVTKVDEPSKYGVVYVMEBEATGRVERFEVKPKIFVGNKINAGTYLL 179
|||||

Db 61 ELIKFKHCHGGEATIMVTKVDEPSKYGVVYVMEBEATGRVERFEVKPKIFVGNKINAGTYLL 120
|||||

Qy 180 NPSVLDRIELRETSIEKEVFPQIADQOQLYAMVLPFGVMDVGPDRYITGLRLYLDLSIRK 239
|||||

Db 121 NPSVLDRIELRETSIEKEVFPQIADQOQLYAMVLPFGVMDVGPDRYITGLRLYLDLSIRK 180
|||||

Qy 240 KSAAKLATGAHVGVNVLVHESAKIGRGCLIGDPDVAIGPGCVWEDGVRLSCTVMRGVRIK 299
|||||

Db 181 KSAARLAAGAHVGVNVLVHESAKIGRGCLIGDPDVAIGPGCVYADGVRLSCTVMRGVRIK 240
|||||

Qy 300 KHACISNSTIGHSTVGOWARIENMTILGEDVHVCDEVYNSGGVLPHPHKEIKSSILKPEI 359
|||||

Db 241 KHACISNSTIGHSTVGOWARIENMTILGEDVHVCDEVYNSGGVLPHPHKEIKSSILKPEI 300
|||||

Qy 360 VM 361
||

Db 301 VM 302
||

RESULT 0
AAG20010
ID AAG20010 standard; Protein; 329 AA.
XX
AC AAG20010;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 22031.
XX
KW protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;

KW	termination sequence.	PR	28-JUN-1999;	99US-0140823.
XX		PR	29-JUN-1999;	99US-0140991.
OS	Arabidopsis thaliana.	PR	30-JUN-1999;	99US-0141287.
XX		PR	01-JUL-1999;	99US-0141842.
PN	EP1033405-A2.	PR	01-JUL-1999;	99US-0142154.
XX		PR	02-JUL-1999;	99US-0142055.
XX		PR	06-JUL-1999;	99US-0142390.
PD	06-SEP-2000.	PR	08-JUL-1999;	99US-0142803.
XX		PR	09-JUL-1999;	99US-0142920.
PF	25-FEB-2000; 2000EP-0301439.	PR	12-JUL-1999;	99US-0142977.
XX		PR	13-JUL-1999;	99US-0143542.
PR	25-FEB-1999;	PR	14-JUL-1999;	99US-0143624.
PR	05-MAR-1999;	PR	15-JUL-1999;	99US-0144005.
PR	09-MAR-1999;	PR	16-JUL-1999;	99US-0144085.
PR	23-MAR-1999;	PR	16-JUL-1999;	99US-0144086.
PR	25-MAR-1999;	PR	19-JUL-1999;	99US-0144325.
PR	29-MAR-1999;	PR	19-JUL-1999;	99US-0144331.
PR	01-APR-1999;	PR	19-JUL-1999;	99US-0144332.
PR	06-APR-1999;	PR	19-JUL-1999;	99US-0144333.
PR	08-APR-1999;	PR	19-JUL-1999;	99US-0144334.
PR	16-APR-1999;	PR	19-JUL-1999;	99US-0144335.
PR	19-APR-1999;	PR	20-JUL-1999;	99US-0144352.
PR	21-APR-1999;	PR	20-JUL-1999;	99US-0144632.
PR	23-APR-1999;	PR	20-JUL-1999;	99US-0144884.
PR	23-APR-1999;	PR	21-JUL-1999;	99US-0144884.
PR	28-APR-1999;	PR	21-JUL-1999;	99US-0145086.
PR	30-APR-1999;	PR	21-JUL-1999;	99US-0145088.
PR	30-APR-1999;	PR	22-JUL-1999;	99US-0145085.
PR	04-MAY-1999;	PR	22-JUL-1999;	99US-0145087.
PR	05-MAY-1999;	PR	22-JUL-1999;	99US-0145089.
PR	06-MAY-1999;	PR	22-JUL-1999;	99US-0145192.
PR	06-MAY-1999;	PR	23-JUL-1999;	99US-0145192.
PR	07-MAY-1999;	PR	23-JUL-1999;	99US-0145218.
PR	11-MAY-1999;	PR	23-JUL-1999;	99US-0145224.
PR	14-MAY-1999;	PR	26-JUL-1999;	99US-0145276.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145913.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145918.
PR	18-MAY-1999;	PR	27-JUL-1999;	99US-0145919.
PR	19-MAY-1999;	PR	28-JUL-1999;	99US-0145951.
PR	20-MAY-1999;	PR	02-AUG-1999;	99US-0146386.
PR	21-MAY-1999;	PR	02-AUG-1999;	99US-0146388.
PR	21-MAY-1999;	PR	02-AUG-1999;	99US-0146389.
PR	24-MAY-1999;	PR	03-AUG-1999;	99US-0147038.
PR	25-MAY-1999;	PR	04-AUG-1999;	99US-0147204.
PR	27-MAY-1999;	PR	04-AUG-1999;	99US-0147302.
PR	28-MAY-1999;	PR	05-AUG-1999;	99US-0147192.
PR	01-JUN-1999;	PR	05-AUG-1999;	99US-0147260.
PR	03-JUN-1999;	PR	06-AUG-1999;	99US-0147303.
PR	04-JUN-1999;	PR	06-AUG-1999;	99US-0147416.
PR	07-JUN-1999;	PR	09-AUG-1999;	99US-0147935.
PR	08-JUN-1999;	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	PR	10-AUG-1999;	99US-0148171.
PR	10-JUN-1999;	PR	11-AUG-1999;	99US-0148319.
PR	14-JUN-1999;	PR	12-AUG-1999;	99US-0148341.
PR	16-JUN-1999;	PR	13-AUG-1999;	99US-0148565.
PR	16-JUN-1999;	PR	13-AUG-1999;	99US-0148684.
PR	17-JUN-1999;	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	PR	27-AUG-1999;	99US-0151065.
PR	21-JUN-1999;	PR	27-AUG-1999;	99US-0151066.
PR	22-JUN-1999;	PR	27-AUG-1999;	99US-0151080.
PR	22-JUN-1999;	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	PR	31-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	PR	07-SEP-1999;	99US-0152363.


```
OY 213 LRFWMVGPQPDYITGLRLYLDSIRKKSAAKATGAHVGVNVLVHESAKIGEGCLIGPD 272
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 181 LRFWMVGPQPDYITGLRLYLDSIRKKSAAKATGAHVGVNVLVHESAKIGEGCLIGPD 240
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
OY 273 VAIGCCVVEGVRLSRCTVGRVRIKKHACISNSITGHSTVGVQWARIENMTILGEDVH 332
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 241 VAIGCCVVEGVRLSRCTVGRVRIKKHACISNSITGHSTVGVQWARIENMTILGEDVH 300
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
OY 333 YCDEYSNGVYVLPKHEKSSILKPEIVM 361
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 301 VSEIYSNGCVVLPKHEKSNILKPEIVM 329
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 10
AAG20011
ID AAG20011 standard; Protein: 302 AA.
XX
AC AAG20011;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 22032.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
FN EP1033405-A2.
XX
DD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125789.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130691.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137724.
PR 07-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 24-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145013.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
```

PR 06-AUG-1999;	99US-0147303.
PR 06-AUG-1999;	99US-0147416.
PR 09-AUG-1999;	99US-0147493.
PR 09-AUG-1999;	99US-0147935.
PR 10-AUG-1999;	99US-0148171.
PR 11-AUG-1999;	99US-0148319.
PR 12-AUG-1999;	99US-0148341.
PR 13-AUG-1999;	99US-0148565.
PR 13-AUG-1999;	99US-0148684.
PR 16-AUG-1999;	99US-0149368.
PR 17-AUG-1999;	99US-0149175.
PR 18-AUG-1999;	99US-0149426.
PR 20-AUG-1999;	99US-0149722.
PR 20-AUG-1999;	99US-0149723.
PR 20-AUG-1999;	99US-0149929.
PR 23-AUG-1999;	99US-0149902.
PR 23-AUG-1999;	99US-0149930.
PR 25-AUG-1999;	99US-0150566.
PR 26-AUG-1999;	99US-0150884.
PR 27-AUG-1999;	99US-0151065.
PR 27-AUG-1999;	99US-0151066.
PR 27-AUG-1999;	99US-0151080.
PR 30-AUG-1999;	99US-0151303.
PR 31-AUG-1999;	99US-0151438.
PR 01-SEP-1999;	99US-0151930.
PR 07-SEP-1999;	99US-0152363.
PR 10-SEP-1999;	99US-0153070.
PR 13-SEP-1999;	99US-0153758.
PR 15-SEP-1999;	99US-0154018.
PR 16-SEP-1999;	99US-0154039.
PR 20-SEP-1999;	99US-0154779.
PR 22-SEP-1999;	99US-0155139.
PR 23-SEP-1999;	99US-0155486.
PR 24-SEP-1999;	99US-0155659.
PR 28-SEP-1999;	99US-0156458.
PR 29-SEP-1999;	99US-0156596.
PR 04-OCT-1999;	99US-0157117.
PR 05-OCT-1999;	99US-0157753.
PR 06-OCT-1999;	99US-0157865.
PR 07-OCT-1999;	99US-0158029.
PR 08-OCT-1999;	99US-0158232.
PR 12-OCT-1999;	99US-0158369.
PR 13-OCT-1999;	99US-0159293.
PR 13-OCT-1999;	99US-0159294.
PR 13-OCT-1999;	99US-0159295.
PR 14-OCT-1999;	99US-0159329.
PR 14-OCT-1999;	99US-0159330.
PR 14-OCT-1999;	99US-0159331.
PR 14-OCT-1999;	99US-0159637.
PR 18-OCT-1999;	99US-0159638.
PR 21-OCT-1999;	99US-0160741.
PR 21-OCT-1999;	99US-0160767.
PR 21-OCT-1999;	99US-0160768.
PR 21-OCT-1999;	99US-0160770.
PR 21-OCT-1999;	99US-0160814.
PR 21-OCT-1999;	99US-0160815.
PR 22-OCT-1999;	99US-0160980.
PR 22-OCT-1999;	99US-0160981.
PR 22-OCT-1999;	99US-0160989.
PR 25-OCT-1999;	99US-0161404.
PR 25-OCT-1999;	99US-0161405.
PR 25-OCT-1999;	99US-0161406.
PR 26-OCT-1999;	99US-0161359.
PR 26-OCT-1999;	99US-0161360.
PR 26-OCT-1999;	99US-0161361.
PR 28-OCT-1999;	99US-0161920.
PR 28-OCT-1999;	99US-0161992.
PR 28-OCT-1999;	99US-0161993.
PR 29-OCT-1999;	99US-0162142.

Query Match 75.2%; Score 1402; DB 21; Length 302;
Best Local Similarity 87.4%; Pred. No. 6.9e-146;

Matches 264; Conservative 21; Mismatches 17; Indels 0; Gaps 0;	
QY 60	MINFLKDFDKLGITITCSQETPLCTAGPLALARKLADGSGQPFVLSNDVISEYPEFA 119
DB 1	MLNFLKDFETKLEIKITCSQETPLCTAGPLALARKLKLDDGSGPEFVLSNDVISEYPLK 60
QY 120	ELIKPHKCHGGEATIMVTKVDEPSKYGVVYMEEATGRVERFEKPKIFVGNKINAGIYLL 179
DB 61	EMLEPHKSHGGEASIMVTKVDEPSKYGVVYMEESTGRVEKFEKPKLYVGNKINAGIYLL 120
QY 180	NPSVLDRIELRPTSIEKEVFPQIAADOOLYAMVLPGFWDVGOPROYITGLRLYLSIRK 239
DB 121	NPSVLDKIELRPTSIEKETFPKIAAAGLYAMVLPGFWDIGOPROYITGLRLYLSLRK 180
QY 240	KSAAKLATGAHVGVNVLVHESAKIGECCLIGPDVAIGPGCVVEDGVRLSRTVMRGVRIK 299
DB 181	KSPAKLTSPHIVGNVLDDETATIGEGCLIGPDVAIGPGCIVESGVRLSRTVMRGVRIK 240
QY 300	KHACISNSIIGHSTVGQWARIENMTILGEDVHVCDEVYSGGVVLPHKEIKSILKPEI 359
DB 241	KHACISSIIGHSTVGQWARIENMTILGEDVHVSDEIYSGGVVLPHKEIKSNILKPEI 300
QY 360	VM 361
DB 301	VM 302
RESULT 11	
AAG46461	
ID	AAG46461 standard; Protein; 302 AA.
XX	AC AAG46461;
XX	DF 18-OCT-2000 (first entry)
XX	DE Arabidopsis thaliana protein fragment SEQ ID NO: 58454.
KW	Protein identification; signal transduction pathway; metabolic pathway;
KW	hybridisation assay; genetic mapping; gene expression control; promoter;
XX	termination sequence.
OS	Arabidopsis thaliana.
PN	EP1033405-A2.
XX	PD 06-SEP-2000.
XX	PF 25-FEB-2000; 2000EP-0301439.
PR 25-FEB-1999;	99US-0121825.
PR 05-MAR-1999;	99US-0123180.
PR 09-MAR-1999;	99US-0123548.
PR 23-MAR-1999;	99US-0125788.
PR 25-MAR-1999;	99US-0126264.
PR 29-MAR-1999;	99US-0126785.
PR 01-APR-1999;	99US-0127462.
PR 06-APR-1999;	99US-0128234.
PR 16-APR-1999;	99US-0128714.
PR 19-APR-1999;	99US-0129845.
PR 21-APR-1999;	99US-0130077.
PR 23-APR-1999;	99US-0130449.
PR 23-APR-1999;	99US-0130510.
PR 23-APR-1999;	99US-0130891.
PR 30-APR-1999;	99US-0131449.
PR 30-APR-1999;	99US-0132048.
PR 04-MAY-1999;	99US-0132407.
PR 05-MAY-1999;	99US-0132484.
PR 06-MAY-1999;	99US-0132485.
PR 07-MAY-1999;	99US-0132486.
PR 11-MAY-1999;	99US-0132487.
PR 14-MAY-1999;	99US-0132863.
PR 14-MAY-1999;	99US-0134256.
PR 14-MAY-1999;	99US-0134218.

PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140333.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140931.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142330.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149920.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.

[illegible]

RESULT 13
AAV79210
ID AAV79210 standard; Protein; 360 AA.
XX
XX
AC AAV79210;
XX
DT 19-JUN-2000 (first entry)
XX
DE Human transferase TRNSFS-2.
XX
KW Transferase; TRNSFS-2; human; antitumour; cancer;
KW gastrointestinal disorder; developmental disorder;
KW genetic disorder; neurological disorder; reproductive disorder;
KW smooth muscle disorder; immunological disorder; inflammation;
KW diagnosis; therapy; mannose-1-phosphate guanyltransferase.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 290
FT Modified-site /note= "potential O-phosphorylation"
FT Modified-site /note= "potential O-phosphorylation"
FT Modified-site 136
FT Modified-site /note= "potential O-phosphorylation"
FT Modified-site 191
FT Modified-site /note= "potential O-phosphorylation"
FT Modified-site 301
FT Modified-site /note= "potential O-phosphorylation"
FT Modified-site 348
FT Modified-site /note= "potential O-phosphorylation"
FT Modified-site 352
FT Modified-site /note= "potential O-phosphorylation"
FT Modified-site 21
FT Modified-site /note= "potential O-phosphorylation"
FT Modified-site 153
FT Modified-site /note= "potential O-phosphorylation"
FT Modified-site 168
FT Modified-site /note= "potential O-phosphorylation"
FT Modified-site 235
FT Modified-site /note= "potential O-phosphorylation"
FT Modified-site 259
FT Modified-site /note= "potential O-phosphorylation"
FT Modified-site 144
FT Modified-site /note= "potential O-phosphorylation"
FT Modified-site 265
FT Modified-site /note= "potential N-glycosylation"
FT Modified-site 271
FT Modified-site /note= "potential N-glycosylation"
FT Modified-site 322
FT Modified-site /note= "potential N-glycosylation"
FT Peptide 256..284
FT Peptide /note= "bacterial hexapeptide transferase signature"
FT Peptide 7..44
FT Peptide /note= "putative ADP-glucose
FT Peptide 106..144
FT Peptide /note= "putative ADP-glucose
FT Peptide /note= "pyrophosphorylase signature"
XX
PN WO200014251-A2.
XX
PD 16-MAR-2000.
XX
PF 09-SEP-1999; 99WO-US20989.
XX
PR 10-SEP-1998; 98US-0150657.
PR 04-NOV-1998; 98US-0186779.
PR 11-MAY-1999; 99us-0133642.
XX

PA (INCY-) INCYTE PHARM INC.
XX Tang YT, Corley NC, Guegler KJ, Baughn MR, Lal P, Yue H;
PI Hillman JL, Azimzai Y;
XX
XX WPI; 2000-256996/22.
DR N-PSDB; AAZ94202.
XX
PT Human transferase proteins useful for preventing, diagnosing and
PT treating cancers and developmental, gastrointestinal, genetic,
PT immunological, neurological, reproductive and smooth muscle disorders -
XX
PS Claim 1; Page 80; 113pp; English.
XX
CC The present sequence is that of human transferase TRNSFS-2, 1 of
CC 15 claimed novel human transferase proteins of the invention (see
CC AAY79209-23)). The sequence was deduced from a cDNA clone (see
CC AAY79209-23)) isolated from an ileum tissue library. It shows homology
CC to mannose-1-phosphate guanyltransferases. TRNSFS-2 is expressed
CC in reproductive and gastrointestinal tissues, especially those
CC associated with cancer and inflammation. The new human transferase
CC proteins and the polynucleotides encoding them can be used in the
CC diagnosis, prevention and treatment of cancer, developmental
CC disorders, gastrointestinal disorders, genetic disorders,
CC immunological disorders, neurological disorders, reproductive
CC disorders, and smooth muscle disorders. The polypeptides can also
CC be used to raise antibodies, and to screen for agonists and
CC antagonists of transferase activity.
XX
SO Sequence 360 AA;
Query Match 63.5%; Score 1183.5; DB 21; Length 360;
Rest Local Similarity 61.2%; Pred. No. 1.3e-121;
Matches 221; Conservative 65; Mismatches 74; Indels 1; Gaps 1;
OY 1 MKALILVGGFGRRLRLPLTLSPKPLVDFANKPMLHIOIEALKEVGTVVLAINYRPEVM 60
DB 1 MKALILVGGYGRRLRLPLTLSTPKPLVDFCNKPKILLHQVEALAAAGVDHVLAVSYMQL 60
OY 61 INFLKDPEDKLGITTCSTOFTEDPLCTAGPLALARDKADCGOPFFVLNSDVISEYFAE 120
DB 61 EKEMKAOEORLGRISMSHEEPLGTAGPLALARDLUSE-TADFFVFLNSDVIICDFPQA 119
OY 121 LKFKHKGHEATIMVTKVDFPSKYGVVYMEATGRVREVERKPFVGNKINAGIYLLN 180
DB 120 MVQFRRHGGESILVTKVEEPSKYGVVCEADTGRHREVERKPFVGNKINAGIYLLN 179
OY 181 PSVLDRIELRPTSTIEKEVFPQIAADQOLYAMVLPGFWMVQGPDRDYITGLRLYLDSTRKK 240
DB 180 PAVLRRIQLQPTSTIEKEVFPIMAKEGLYAMELQGFWMIDIGQPKDFLTGCLFLQSLRQK 239
OY 241 SAAKLATGAHVGVGNVLVHESAKIGEGCLIGPDVAIGPCVVEDGVRSLRCTVMRGVRIKK 300
DB 240 QPERLCGPGIYGNVLVDFPSARIGQNCISGPNVSLGPGVYVEDGVCIRRCTVLRDAKRS 299
OY 301 HACISNSIIGHSTVGQWARIENNTILGEDVHVCDEVSYNGVVLPHKETSILKPEIV 360
DB 300 HSWLESCIVGWRRCRGQWVRMENVTVLGEDVIYNDELYLNGASVLPKHSIGESVPEPRII 359
OY 361 M 361
DB 360 M 360
RESULT 14
AAV52357
ID AAV52357 standard; Protein; 363 AA.
XX
XX AAV52357;
XX
XX 25-JAN-2002 (first entry)
XX
DE NDP-hexose pyrophosphorylase homolog.

XX Geminivirus; plant; viral infection; transgenic plant;
KW tomato yellow leaf curl virus.
XX Schizosaccharomyces pombe.
XX FR2806095-A1.
XX 14-SEP-2001.
XX 10-MAR-2000; 2000FR-00031140.
XX 10-MAR-2000; 2000FR-00031140.
XX (GENT-) GENTECH SARL.
XX Bejarano ER, Castillo GA, Collinet D, Donoso CI, Iniesta JR;
PI Grevesse C, Hericourt F;
XX WPI; 2001-628275/73.
DR N-PSDB; ABA01227.
XX New polynucleotides for producing transgenic plants resistant to
PT geminivirus infection comprising polynucleotides encoding proteins
PT which interact with at least one of the products of the geminivirus
PT genome -
XX Claim 4; Pages 38-39; 106pp; French.
XX The present invention relates to coding sequences encoding proteins which
CC interact with at least one of the six products of the geminivirus genome
CC necessary for infection of a plant by the virus. The present sequence is
CC one such protein. The coding sequences are useful for producing
CC transgenic plants resistant to geminivirus infection, particularly tomato
CC yellow leaf curl virus.
XX Sequence 363 AA;
SQ
Query Match 63.0%; Score 1175; DB 22; Length 363;
Best Local Similarity 62.1%; Pred. No. 1.1e-120;
Matches 226; Conservative 54; Mismatches 80; Indels 4; Gaps 4;
QY 1 MKALILVGGFCTRLRLPTLSPFKPLVDFAFKPMILHQLKEVGTWVLAINRPEVM 60
DB 1 MKALILVGGFCTRLRLPTLTPKPLVEFGNPKMILHQLVEALAAAGTVDIVLAVNRP 60
QY 61 INFLKDFEDKLGITITCSQETPLGTAGPLALARDKLDAGSGQPFVFLNSDVISEYPAE 120
DB 61 VEALKYKEYNVNTFSEVNEPLGTAGPLALARDILAK-DHSPPFVFLNSDVISEYPAE 119
QY 121 LIKFKHCHGEATIMVTKVDEPSKYGVV-MEEATGRVERVEKPKIFVGNKINAGIYLL 179
DB 120 LAAPHKAGAGGTIVTKVEEPSKYGVVHVHPNSSESLIERVEKPEVFNRRINGGIYIL 179
QY 180 NPSVLDRIELRPTSTIEKEVFQIAADQOLYAMVLPFGWMDYQPRDYITGLRLYLDLSIRK 239
DB 180 NPSVLDRIELRPTSTIEKEVFQIAADQOLYAMVLPFGWMDYQPRDYITGLRLYLDLSIRK 239
QY 240 KSAAKLA-TGAHVGNVLVHESAKIGEGCLIGDPDAIGPGCVVEDGVRLSRCTVMRGVRI 298
DB 240 HKPEILAPASSNIIGNVLIDPSATIGNCKIGPNVIGPNVTIGDGVRLQRCAILKSSRV 299
QY 299 KKHACINSIIGWSTVGQWARIENMVLIGDGVHVCDEVYNGGVLVPHKEIKSILKP- 357
DB 300 RDHAWKSSIVGNSTLGSRLNVSVLGDVVVNDIYVNGGSIILPHKSIANSIEVP 359
QY 358 EIVM 361
DB 360 TIVM 363
RESULT 15
ABB57869

ID ABB57869 standard; Protein; 369 AA.
XX ABB57869;
XX 26-MAR-2002 (first entry)
XX Drosophila melanogaster polypeptide SEQ ID NO 399.
DE Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX Drosophila melanogaster.
XX WO200171042-A2.
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US09231.
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX (PEKE) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
PI WPI; 2001-656860/75.
DR N-PSDB; ABL01972.
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX Disclosure; SEQ ID NO 399; 21pp + Sequence Listing; English..
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB2072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pot_sequences.
XX Sequence 369 AA;
SQ
Query Match 62.9%; Score 1173; DB 22; Length 369;
Best Local Similarity 61.7%; Pred. No. 1.9e-120;
Matches 222; Conservative 64; Mismatches 72; Indels 2; Gaps 2;
QY 2 KALILVGGFCTRLRLPTLSPFKPLVDFAFKPMILHQLKEVGTWVLAINRPEVM 61
DB 12 RALILVGGFCTRLRLPTLSTPKPLVEFAFKPMILHQLVEALVADAGCROVILAVSRQME 71
QY 62 NLFKDFEDKLGITITCSQETPLGTAGPLALARDKLDAGSGQPFVFLNSDVISEYPAE 121
DB 72 KELKVEAKLGLVELIFSHETPLGTAGPLALAKTILA-ASSEPPFVFLNSDVICDFPKQL 130
QY 122 IFHKHCHGEATIMVTKVDEPSKYGVVVEEATGRVERVEKPKIFVGNKINAGIYLLNP 181
DB 131 VQFHCNKGEGTIVTKVEEPSKYGVVLYDE-NGCKNFTEKPEQEFVSNKINAGIYIFNP 189
QY 182 SVLDRIELRPTSTIEKEVFQIAADQOLYAMVLPFGWMDYQPRDYITGLRLYLDLSIRKS 241
DB 190 SVLDRIELRPTSTIEKEVFQIAADQOLYAMVLPFGWMDYQPRDYITGLRLYLDLSIRKS 249
QY 242 AAKLATGAHVGNVLVHESAKIGEGCLIGDPDAIGPGCVVEDGVRLSRCTVMRGVRIKKH 301
DB 250 SPKLYTGPVGNVNLVDPTAKIGEGCRIGPNVTIGDPVDVIEDGVCIKRSILKAGIYVRSH 309

Search completed: November 27, 2002, 05:07:04
Job time : 70 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 27, 2002, 05:07:12 ; Search time 26 Seconds
(without alignments)
408.526 Million cell updates/sec

Title: US-09-374-967-2
Perfect score: 1864
Sequence: 1 MKALLVGGFGRRLRLTL.....GVLPHPKEIKSSILKPEIVM 361

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues
Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/iaa/5A-COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B-COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A-COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B-COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS-COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					Description	
Result No.	Score	% Match	Length	ID	Description	
1	1178	63.2	361	3	US-09-032-372-12	Sequence 12, Appl
2	401	21.5	421	3	US-09-032-372-3	Sequence 3, Appli
3	307	16.5	355	1	US-08-196-218-34	Sequence 34, Appl
4	307	16.5	355	1	US-08-681-953-34	Sequence 34, Appl
5	296	15.9	355	4	US-09-194-905-10	Sequence 10, Appl
6	241.5	13.0	461	4	US-09-134-001C-3316	Sequence 3316, Ap
7	239	12.8	146	4	US-09-370-838-190	Sequence 190, App
8	228.5	12.3	224	4	US-09-199-637A-175	Sequence 175, App
9	193	10.4	431	1	US-08-469-202-14	Sequence 14, Appl
10	193	10.4	431	2	US-08-484-434C-14	Sequence 14, Appl
11	191	10.2	431	1	US-08-090-523-4	Sequence 4, Appli
12	191	10.2	431	1	US-08-398-627-4	Sequence 4, Appli
13	191	10.2	431	1	US-08-406-858-4	Sequence 4, Appli
14	191	10.2	431	5	PCT-US91-04036-4	Sequence 4, Appli
15	191	10.2	431	5	PCT-US91-05275-4	Sequence 4, Appli
16	191	10.2	488	4	US-09-444-728-2	Sequence 2, Appli
17	190	10.2	292	3	US-09-320-878-14	Sequence 14, Appl
18	190	10.2	292	4	US-09-105-537-12	Sequence 12, Appl
19	190	10.2	3782	4	US-09-105-537-4	Sequence 4, Appli
20	188	10.1	431	1	US-08-090-523-2	Sequence 2, Appli
21	188	10.1	431	1	US-08-398-627-2	Sequence 2, Appli
22	188	10.1	431	1	US-08-406-858-2	Sequence 2, Appli
23	188	10.1	431	5	PCT-US91-04036-2	Sequence 2, Appli
24	188	10.1	431	5	PCT-US91-05275-2	Sequence 2, Appli
25	186.5	10.0	518	1	US-08-485-241-3	Sequence 3, Appli
26	186.5	10.0	518	2	US-08-874-162-3	Sequence 3, Appli
27	186	10.0	517	1	US-08-485-241-5	Sequence 5, Appli

28	186	10.0	517	2	US-08-874-162-5	Sequence 5, Appli
29	186	10.0	518	1	US-08-299-675-2	Sequence 2, Appli
30	179.5	9.6	229	3	US-09-024-023-2	Sequence 2, Appli
31	179.5	9.6	229	4	US-09-531-111-2	Sequence 2, Appli
32	174	9.3	459	3	US-08-971-782-2	Sequence 2, Appli
33	174	9.3	459	4	US-09-309-026-2	Sequence 2, Appli
34	169.5	9.1	471	2	US-08-535-276-4	Sequence 4, Appli
35	169.5	9.1	471	4	US-09-335-234-4	Sequence 4, Appli
36	168.5	9.0	295	4	US-09-134-001C-4466	Sequence 4466, Ap
37	166.5	8.9	507	1	US-08-097-829-6	Sequence 6, Appli
38	166.5	8.9	507	1	US-08-577-403-6	Sequence 6, Appli
39	163	8.7	95	3	US-09-024-023-4	Sequence 4, Appli
40	163	8.7	95	4	US-09-531-111-4	Sequence 4, Appli
41	163	8.7	521	1	US-08-090-523-8	Sequence 8, Appli
42	163	8.7	521	1	US-08-398-627-8	Sequence 8, Appli
43	163	8.7	521	1	US-08-406-858-8	Sequence 8, Appli
44	163	8.7	521	5	PCT-US91-04036-8	Sequence 8, Appli
45	163	8.7	521	5	PCT-US94-05275-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-09-032-372-12
; Sequence 12, Application US/09032372
; Patent No. 6008337
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Yue, Henry
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: CELL CYCLE RELATED PROTEINS
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/032,372
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0478 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 361 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1292898
; US-09-032-372-12

Query Match 63.2%; Score 1178; DB 3; Length 361;
Best Local Similarity 61.0%; Pred. NO. 6.5e-123;
Matches 221; Conservative 56; Mismatches 83; Indels 2; Gaps 2;

QY 1 MKALILVGGGTRLRPLTLSPKPLVDFANKPMILHIOEALKEVGVTEVVLAINVRP 60
DB 1 MKGLILVGGGTRLRPLTLSPKPLVDFANKPMILHIOEALKEVGVTEVVLAINVRP 60

QY 61 INFLKDFEDKLGITITCSOETEPICAGPLALADKLADGSGQFFVNLSDVISEYFAE 120
DB 61 VETLKKEKEYGNYNITFSVETEPICAGPLALADKLADGSGQFFVNLSDVISEYFAE 119

QY 121 LIKPHKHCHGEATIMTKVDPSKYGVVMEET-GRVERVEKPKIFVGNKINAGIYLL 179
DB 121 LIKPHKHCHGEATIMTKVDPSKYGVVMEET-GRVERVEKPKIFVGNKINAGIYLL 179

QY 120 LADFHAKHGKGTIVATKVDPSKYGVVMEET-GRVERVEKPKIFVGNKINAGIYLL 179
DB 120 LADFHAKHGKGTIVATKVDPSKYGVVMEET-GRVERVEKPKIFVGNKINAGIYLL 179

QY 180 NPSVLDRIELRPPSIEKEVEFPQTAADQOLYAMVLPFGVMDVGPDRDITGLRLYLSIRK 239
DB 180 NPSVLDRIELRPPSIEKEVEFPQTAADQOLYAMVLPFGVMDVGPDRDITGLRLYLSIRK 239

QY 240 KSAAKLATGARHVGNVLVHESAKIGECCLICPDVAIGPGCVVEDGVRLSRTVMRGVRIK 299
DB 240 KSAAKLATGARHVGNVLVHESAKIGECCLICPDVAIGPGCVVEDGVRLSRTVMRGVRIK 299

QY 300 KHACISNITWSTVGOMARIENMTILGEDVHVCDVYSGVGVLPKHKEIKSSILRPEI 359
DB 300 KHACISNITWSTVGOMARIENMTILGEDVHVCDVYSGVGVLPKHKEIKSSILRPEI 359

QY 360 VM 361
DB 360 IM 361

RESULT 2

US-09-032-372-3
Sequence 3, Application US/09032372
Patent No. 6008337

GENERAL INFORMATION:

APPLICANT: Bandman, Olga
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
APPLICANT: Guegler, Karl J.
APPLICANT: Yue, Henry
APPLICANT: Lal, Preeti
TITLE OF INVENTION: CELL CYCLE RELATED PROTEINS
NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/032,372
FILING DATE: Herewith

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0478 US

TELEPHONE: 650-855-0555

TELEFAX: 650-845-4166

TELEX:

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 421 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: COLANOT23
CLONE: 1693222
US-09-032-372-3

Query Match 21.5%; Score 401; DB 3; Length 421;

Best Local Similarity 29.1%; Pred. NO. 4.5e-36;

Matches 123; Conservative 61; Mismatches 159; Indels 80; Gaps 13;

QY 1 MKALILVGG--FGRRLRPLTLSPKPLVDFANKPMILHIOEALKEV-GVTEVVLAINVRP 57
DB 2 LKAVILCGPKQKGRFRLSPKPLVDFANKPMILHIOEALKEV-GVTEVVLAINVRP 61

QY 58 -EVMINFLKDFEDKLGITITCSOETEPICAGPLALADKLADGSGQFFVNLSDVISEY 116
DB 62 DEPLTQFLFAAQEENLFRVYLQEFAPLGTGGGLYHFRQILLAGSPEAFVNLADVCSD 121

QY 117 PFAELIKPHKHCHGEATIMTKVD--EPSKYGVVMEET-GRVERVEKPKIFVGNKIN 174
DB 122 PLSAMLEAHRORHPFLLLGTANTQSIYCCIVENQTHVHVEKPSFISDIINC 181

QY 175 GYLLNPSVL-----DRIELRP-----TSIEKEVFPQIAADQOLYAMV 213
DB 182 GYLLNPSVL-----DRIELRP-----TSIEKEVFPQIAADQOLYAMV 213

QY 214 PGFWMVGPDRDITGLRLYLSIRKSAKIGECCLICPDVAIGPGCVVEDGVRLSRTVM 268
DB 242 DGIRSQKISAGSALYASRLYSLRYQDTHPERLA--KHTPGGMDPRECYHPPDRG-- 295

QY 269 IGPDAVIGP-----GCWVEDGVRLSRL--TVMRGVRIKHKACISNSTIGHSRVG 316
DB 296 -GPLGCGAPQRLHREGGDRGRCVCGSSGRASSMEPLCRSTRV---FCIASWAGGAPWD 351

QY 317 QWARIENMTI-----LGEDVHVCDVYSGVGVLPKHKEIKSSILRPEI 350
DB 352 RAWRVPPVTLPTTPEPAWTVRASSRTGSCCLLSPSMGCRVRIPAEVLILNSIVLPKH 411

QY 351 KSS 353
DB 412 SRS 414

RESULT 3

US-08-196-218-34
Sequence 34, Application US/08196218
Patent No. 5614619

GENERAL INFORMATION:

APPLICANT: Pieperberg, Wolfgang
APPLICANT: Stockmann, Michael
APPLICANT: Taleghani, Kamiz Mansouri
APPLICANT: Distler, Jurgen
APPLICANT: Grabley, Susanne
APPLICANT: Sichel, Petra
APPLICANT: Brau, Barbara
TITLE OF INVENTION: Secondary-Metabolite Biosynthesis Genes
TITLE OF INVENTION: FROM ACTINOMYCETES, METHOD OF ISOLATING THEM, AND THEIR
NUMBER OF SEQUENCES: 34
TITLE OF INVENTION: Use.

CORRESPONDENCE ADDRESS:

ADDRESSEE: Finnegan, Henderson, Parabow, Garrett &
ADDRESS: Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: United States
ZIP: 20005-3315

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/196,218
FILING DATE: 25-AUG-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ogden, Stasia L.
REGISTRATION NUMBER: 36,228
REFERENCE/DOCKET NUMBER: 02481.1372-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 355 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-196-218-34

Query Match 16.5%; Score 307; DB 1; Length 355;
Best Local Similarity 28.4%; Pred. No. 1.1e-25;
Matches 100; Conservative 59; Mismatches 137; Indels 56; Gaps 12;
QY 1 MKALILVGGFGTRRLPTLSFPKPLVDFAFKPMILHQIEALKEVGVTEV-VLAINYRPEV 59
Db 1 MKALVLAGGSGTRLRPFSSMPKQLPIANTPVLVHVLNAVRELGVTEGVVGNRGPEI 60
QY 60 MINFLKDFEDKLG-----ITTCSETEPLGTAGPLALARDKLADGSGQFFVLNSDVI 113
Db 61 -----EAVLDGARFDVRIITYIQDAPRGLAHTVSIARGFLGD---DDFVMYLGDNM 109
QY 114 SEYPPAELIKFKHKGGEATIMVTKVDEPSKYGVVVMEEATGRVERFEKPKIFVGNKIN 173
Db 110 LPDGVTEIAEEFTQRPAQVVKVPPDRSFGVAEL-GPDGEVLRLVEKWPQPSDMAL 168
QY 174 AGIYLLNPSVLDRI-ELRPTS-----IEKEVFPQIAADQQLYAMVLPFGFWMVDGQPRDYI 227
Db 169 IGVIYFETAIIHQAVAAISPSSRGELEITDAVQWLVTSGADVRSALYDGYWKDGTGRVEDVL 228
QY 228 TGLRLYLDLSIRKSAKALATGAHVGNVNLVHESAK-----IGEGCL-----IGP 271
Db 229 ECNSHLLDGLTPRDVGQDADSVLVGRVYIEAGARIVRSRVEGPAIIGAGTVLQDSQVGP 288
QY 272 DVAIGPGCVVED-----GVRLSRTV--MRGVRIKKHACISNIIIGHSTVG 316
Db 289 HTSIGRDTCTVDSRLSGSIALDEASVTGVRGLR-----NSLIGRAASVG 332

RESULT 4
US-08-681-953-34
Sequence 34, Application US/08681953
Patent No. 5710032
GENERAL INFORMATION:
APPLICANT: Piepersberg, Wolfgang
APPLICANT: Stockmann, Michael
APPLICANT: Taleghani, Kamriz Mansouri
APPLICANT: Distler, Jurgen
APPLICANT: Grabley, Susanne
APPLICANT: Sichel, Petra
APPLICANT: Brau, Barbara
TITLE OF INVENTION: Secondary-Metabolite Biosynthesis Genes
TITLE OF INVENTION: From Actinomycetes, Method of Isolating Them, and Their
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner
STREET: 1300 I Street, N.W.

CITY: Washington
STATE: D.C.
COUNTRY: United States
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/681,953
FILING DATE: 30-JUL-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/196,218
FILING DATE: 25-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ogden, Stasia L.
REGISTRATION NUMBER: 36,228
REFERENCE/DOCKET NUMBER: 02481.1372-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 355 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-681-953-34

Query Match 16.5%; Score 307; DB 1; Length 355;
Best Local Similarity 28.4%; Pred. No. 1.1e-25;
Matches 100; Conservative 59; Mismatches 137; Indels 56; Gaps 12;
QY 1 MKALILVGGFGTRRLPTLSFPKPLVDFAFKPMILHQIEALKEVGVTEV-VLAINYRPEV 59
Db 1 MKALVLAGGSGTRLRPFSSMPKQLPIANTPVLVHVLNAVRELGVTEGVVGNRGPEI 60
QY 60 MINFLKDFEDKLG-----ITTCSETEPLGTAGPLALARDKLADGSGQFFVLNSDVI 113
Db 61 -----EAVLDGARFDVRIITYIQDAPRGLAHTVSIARGFLGD---DDFVMYLGDNM 109
QY 114 SEYPPAELIKFKHKGGEATIMVTKVDEPSKYGVVVMEEATGRVERFEKPKIFVGNKIN 173
Db 110 LPDGVTEIAEEFTQRPAQVVKVPPDRSFGVAEL-GPDGEVLRLVEKWPQPSDMAL 168
QY 174 AGIYLLNPSVLDRI-ELRPTS-----IEKEVFPQIAADQQLYAMVLPFGFWMVDGQPRDYI 227
Db 169 IGVIYFETAIIHQAVAAISPSSRGELEITDAVQWLVTSGADVRSALYDGYWKDGTGRVEDVL 228
QY 228 TGLRLYLDLSIRKSAKALATGAHVGNVNLVHESAK-----IGEGCL-----IGP 271
Db 229 ECNSHLLDGLTPRDVGQDADSVLVGRVYIEAGARIVRSRVEGPAIIGAGTVLQDSQVGP 288
QY 272 DVAIGPGCVVED-----GVRLSRTV--MRGVRIKKHACISNIIIGHSTVG 316
Db 289 HTSIGRDTCTVDSRLSGSIALDEASVTGVRGLR-----NSLIGRAASVG 332

RESULT 5
US-09-194-905-10
Sequence 10, Application US/09194905
Patent No. 6306627
GENERAL INFORMATION:
APPLICANT: Decker, Heinrich
TITLE OF INVENTION: ISOLATION OF THE BIOSYNTHESIS GENES FOR
TITLE OF INVENTION: PSEUDO-OLIGOSACCHARIDES FROM STREPTOMYCES GLAUCESCENS
TITLE OF INVENTION: GLA.O AND THEIR USE
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY & LARDNER
STREET: 3000 K Street, N.W.

CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/194,905
FILING DATE: 20-JUL-1998
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/EP97/02826
FILING DATE: 30-MAY-1997

PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 19622783.6
FILING DATE: 07-JUN-1996

ATTORNEY/AGENT INFORMATION:
NAME: Granados, Patricia D.
REGISTRATION NUMBER: 33,683

REFERENCE/DOCKET NUMBER: 026083/0193

TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300

TELEFAX: (202) 672-5399

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 355 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-194-905-10

Query Match 15.9%; Score 296; DB 4; Length 355;

Best Local Similarity 27.5%; Pred. No. 1.8e-34;

Matches 95; Conservative 58; Mismatches 160; Indels 32; Gaps 9;

QY 1 MKALILVGGFGTPLRDLTSLSPKPLVDFAFKPMILHQTALKEKVGVTVEVLATN-VRPEV 59
DB 1 VIALVLAGGTGSKURPTHTAAKLPVLIANKPVLFTALESIAAGVREAGVVGAYGRE- 59
QY 60 MIFLDFEDKLGITITCSQETPLTAGPLALARKLADGSGQPFVYLSNDVISEYFFA 119
DB 60 -IRELTGDCGTAFGLRTYHLHOPRLGLAHAVRIARGFLCD--DDFLYLGDNVLPQGV 115
QY 120 ELIKFKHCHGGEATIMVTKVDEPSKYGVVMEERATGRVERFEKPKTFVGNKINAGIYLL 179
DB 116 DPARQSAADPAARLLLTVPADPSAFGEAEV-DADGNVLRLEEKPDVFRSGLALIGVYAF 174
QY 180 NPSVLDRIE-LRPTS-----IEKEVFPDIAADOLYAMVLPGFMDVQCPDRDVTCLRLY 233
DB 175 SPVHAEVRAITPSARGELEITHAVOMMDRGURVRAETTRPWRDTPGSAEDMLEVNRHV 234
QY 234 LDISIRKSAK-----LATGAHVGNVNLVHESAKTGCGLGPDVAIG 276
DB 235 LDGLEIRIEGKVDAHSLTVGRVRAECAIVRGSHVGPVIGACVAVNSN-VGPTYSIG 293
QY 277 PCGVYEDGVRSLRCTVMRGVRIKKHACISNLIHGHSTYGVQWARI 321
DB 294 EDCRVDSA-IEVSVLLRGAQVEGASRIEASLIRGAVGVPAPRL 337

RESULT 6

US-09-134-001C-3316

Sequence 3316, Application US/09134001C

Patent No. 6380370

GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: GTC-007

CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-09
PRIOR APPLICATION NUMBER: US 60/035,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 3316
LENGTH: 461
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3316

Query Match 13.0%; Score 241.5; DB 4; Length 461;

Best Local Similarity 25.7%; Pred. No. 3.3e-18;

Matches 98; Conservative 66; Mismatches 159; Indels 59; Gaps 15;

QY 3 ALILVGGFGTPLRDLTSLSPKPLVDFAFKPMILHQTALKEKVGVTVEVLAINVRPEVMIN 62
DB 15 AIIILAAAGGTMMKSKY---KVLHEVAGKPMVHVNKNVKGAGVDQIVIIIGHGAE---- 67
QY 63 FLKDFEDKLGITITCSQETPLTAGPLALARKLADGSGQPFVYLSNDVISEYFFAEL 121
DB 68 ---SVKDTLGNOSLYSPQDKOLCTAHAVKMAHEHLADKECTILVVGCDPLITVOTLOS 124
QY 122 IKFKHCHGGEATIMVTKVDEPSKYGVVMEERATGRVERFEKPKIFVGNK-----INAGIY 177
DB 125 IEHHESTQSHVTVLSASTINPYGRIIRNH-NGILERIVEEKDANDSEAIKBISSGIF 183
QY 178 LLNPSVL-DRIELRPTSTIEKEVFPDIAADOLYAMVLPGFMDVQCPDRDVT-----G 229
DB 184 AFNRYLFLEKLE-----QVKNDAQGEYLLDFVLSLILKDGKAKVYCTEDFDEIIG 235
QY 230 L--RLYLDSTIRKSAK-----LATGAHV--GNVLVHESAKTGCGLGPDVAIGPCV 280
DB 236 VNDRLMLSEAEKALOORINRYHMEGVTIIDPSSTFICTDVKIGIDTIEPCVRIGGHTT 295
QY 281 VEDGY-----RLSRCTVMRGVRIKKHACISNLIHGHSTYGVQWARIENMTILGEDVHVC 334
DB 296 IEEDVDTGQVSEINNSTHSHNATKO-SVTNDSIVGENTTVCPAOLRPGSNLASEVKV- 353
QY 335 DEVYSGGVVLPHPHEIKSSILK 356
DB 354 -----GNFV----EYKKAADIK 365

RESULT 7

US-09-370-838-190

Sequence 190, Application US/09370838

Patent No. 6444425

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.

APPLICANT: Lodes, Michael J.

APPLICANT: Mohamath, Radooh

APPLICANT: Secrist, Heather

TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF

TITLE OF INVENTION: LUNG CANCER AND METHODS FOR THEIR USE

FILE REFERENCE: 210121.475C1

CURRENT APPLICATION NUMBER: US/09/370,838

CURRENT FILING DATE: 1999-08-09

EARLIER APPLICATION NUMBER: US 09/285,323

EARLIER FILING DATE: 1999-04-02

NUMBER OF SEQ ID NOS: 289

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 190

LENGTH: 146

TYPE: PRT

ORGANISM: Homo sapien

US-09-370-838-190

Query Match

Best Local Similarity 12.8%; Score 239; DB 4; Length 146;

Matches 47; Conservative 23; Mismatches 33; Indels 28; Gaps 1;

QY	251	VWGNLVHESAKTGEGLIGPDVAIGPGCVVEDGVRLSRCTVMRGVRIKKHACISNIITG	310
Db	9	LLGNVYIHPHTAKVAPSAVLGPNVSIICKGVTYVGGVRLRESIVLHGATIQEHTCVLHSIVG	68
QY	311	WHSTVGVQWARIEN-----WTILGEDVHVHCDVYVSNNG	342
Db	69	WGSTVGRWARVEGTPSPDNPNDPRARMDSELSFKDGKLLPATILGCRVRIAPAEVLILNS	128
QY	343	VVLPHKEIKSS	353
Db	129	IVLPHKELSR	139

```

RESULT 8
US-09-199-637A-175
; Sequence 175, Application US/09199637A
; Patent No. 6355411
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick
; APPLICANT: Goodman, Howard M.
; APPLICANT: Rahme, Laurence G.
; APPLICANT: Mahajan-Miklos, Shalina
; APPLICANT: Tan, Man-Wah
; APPLICANT: Cao, Hui
; APPLICANT: Drenkard, Eliana
; APPLICANT: Tsongalis, John
; TITLE OF INVENTION: SEQUENCES-ASSOCIATED
; TITLE OF INVENTION: SEQUENCES AND USES TH
; FILE REFERENCE: 00786/361002
; CURRENT APPLICATION NUMBER: US/09/199,637A
; CURRENT FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 60/066,517
; PRIOR FILING DATE: 1997-11-25
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 175
; LENGTH: 224
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-199-637A-175

```

Query Match	12.3%	Score	228.5;	DB	4;	Length	224;
Best Local Similarity	33.3%;	Pred. No.	2.9e-17;				
Matches	75;	Conservative	39;	Mismatches	90;	Indels	21; Gaps
							8;
QY	1	MMKALILVGGFGTRLRPTLTLGFPKKPLVDVFANKPMTLHQIEALKVEGVTEVVLAINVRPEVM	60				
		: : : : : : : : : : : : :					
Db	1	MMKAMILAAGRGERNRPITLHTPKPLIEAAGVPLTERQLLARQAGVDDWVINHAWLGEQI	60				
QY	61	INFLKDPEDKLGITITCSQETPELTGTAPITALARDKIADSGSQPFVVNSDVISEYPFAE	120				
		: : : : : : : : : : : : : :					
Db	61	EAYLGD--GSRLGGRIAYSPEEPLETGGIFRALPLLEGE---QPFLINGDVSDFDYSR	116				
QY	121	LIRFKHKGCGEATIMVKTVDEPSKYGVVME-EATGRVERFEVKPIFVGKNKINAGIVLL	179				
		: : : : : : : : : : : : : :					
Db	117	L---HLADGLAHULV--VDNPAHHDPAGDHLDAGRVTGETRAG-----GNLTYSYGIAVL	167				
QY	180	NPSVLDRIELRPTSIEKEVPFQIAADQ---OLYAMVLPGFWMDYG	221				
		: : : : : : : : : : : : : :					
Db	168	HPALFESCQGAFLKLPALLRKAIAGAARVSGSEHYR----GOWVDYG	208				
		: : : : : : : : : : : : : :					

RESULT 9
US-08-469-202-14
; Sequence 14, Application US/08469202
; Patent No. 5750875
; GENERAL INFORMATION:
; APPLICANT: STALKER, DAVID
; TITLE OF INVENTION: GLYCOGEN BIOSYNTHETIC ENZYMES IN PLANTS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CALGENE, INC.

STREET: 1920 FIFTH STREET
CITY: DAVIS
STATE: CA
COUNTRY: USA
ZIP: 95616
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.1
SOFTWARE: Microsoft Word 5.1 (a)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,202
FILING DATE: 6-JUNE-95
CLASSIFICATION: 800
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/016,881
FILING DATE: 11-FEB-1993
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Elizabeth Lassen
REGISTRATION NUMBER: 31,845
NAME: Donna E. Scherer
REGISTRATION NUMBER: 34,719
NAME: Carl J. Schwedler
REGISTRATION NUMBER: 36,924
REFERENCE/DOCKET NUMBER: CGNE 93-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 916-753-6313
TELEFAX: 916-753-1510
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 431 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-469-202-14

Query Match	10.4%	Score 193;	DB 1;	Length 431;
Best Local Similarity	22.3%	Pred. No. 7.7e-13;		
Matches 94;	Conservative	61;	Mismatches 170;	Indels 98;
				Gaps 15;

QY	3	ALLIVGGFTRLRPLRLSLFPKPLDVFANKPMIL-HQTEALKEVGVTEVTLAINRPEVMI	61
DB	22	ALLIAGGTRLRKDLNKRKAPAVHFGGKFRIIDFALSNCINSIRRMGVITQYQSHTLV	81
QY	62	NFLK-----DPEDKLGITITCSQTEPLGTAGPLALARDKLDAGSGOPFFVLNS	110
DB	82	QHTQRCGSFFNEEMNEFVDLLPAQORMKGENWYRGTAADVTQNLDIRRYKAEVVVILAG	141
QY	111	DVTSEYFPFABELIKPHKCHGGEATI--MVTKVDEPSKYGVVMEEATGRVSRFVEKKPIFV	168
DB	142	DHIYKODYSRMLIDHVEKGARCTVACMPVPTEESAFGVMAVDE-NDKIIIEFVEKP----	196
QY	169	GNK-----INAGIYLLNPSVL-----DRIELRPTSTIEKEVFPQIAADQOOLYA	210
DB	197	ANPSPMPNDPSKSLASMGIVYFDADYLYELLEEDDRDENSHPDKLIPKITEAGLAYA	256
QY	211	MVLP-----GFWDMVGOPRDYITG-----LRLYLDSTRKK	240
DB	257	HPFPLSCVQSDPDPAEPYWRDVGTLWAYKANLDLASVVPPELDMYDRNWPRTYNESLPPA	316
QY	241	SAAKLATGAHVGNVLVHESAKTGEGLIGPDVAIGPCGVVEDGVRLSRCTVVRGVRIKK	300
DB	317	KFYQDRSGSH--GMTL---NSLVSDDGCVIS-----GSVVQSVLFSR-----VRVNS	358
QY	301	HACISNIIIGHWSTVGOWARIENNTI-----LGEDVHVCDBVY--SNGGVVLP	347
DB	359	FCDDISANVLLPEVWVGSRCLRCVIDRACVIEPGMWIGENAEADARRFFRSEGEIVLVT	418
QY	348	KEI 350	
DB	419	REM 421	

1

Db 197 ANPPMPNDPSKSLASMGIVYFDADYLYELLEDDRRDENSNDHDFGKDLIPKITEAGLAYA 256
QY 211 MVLP-----GFMDVQPRDYITG-----LRLYDSIRKK 240
Db 257 HPFPLSCVQSDPDAPYWRDVGTLAYWKANLDLASVVPPELMDYRNPRTYNESLPPA 316
QY 241 SAAKATGAHVGNVNLVHESAKIGEGCLIGPDVAIGPCVVEDGVRLSCTVMRGVRIKK 300
Db 317 KFDQDRSGH--GMTL---NSLVSDGCVIS-----GSVVQSVLFSR-----VRVNS 358
QY 301 HACISNSIIGHSTVGQWARIENMTI-----LGEDVHVCDEVY--SNGGVVLPH 347
Db 359 FCNIDSALLPEVWVGRSCLRRRCVDRACVPIEPMGVIGENAEEDARRFYRSEEGIVLVT 418
QY 348 KEI 350
Db 419 REM 421

RESULT 12
US-08-398-627-4
; Sequence 4, Application US/08398627
; Patent No. 5608149
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Stark, David M.
; TITLE OF INVENTION: Enhanced Starch Biosynthesis
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Grace L. Bonner, Monsanto Co. BB4F
; STREET: 700 Chesterfield Parkway No. 5608149th
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/398,627
; FILING DATE: 03-MAR-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/090,523
; FILING DATE: 12-JUL-1993
; APPLICATION NUMBER: US 07/709663
; FILING DATE: 07-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/53763
; FILING DATE: 18-JUN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Bonner, Grace L.
; REGISTRATION NUMBER: 32,963
; REFERENCE/DOCKET NUMBER: 38-21(10559)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314) 537-7286
; TELEFAX: (314) 537-6047
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 431 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-398-627-4

Query Match 10.2%; Score 191; DB 1; Length 431;
Best local similarity 22.5%; Pred. No. 1.3e-12;
Matches 95; Conservative 60; Mismatches 170; Indels 98; Gaps 15;

QY 3 ALLIVGGFGLRLPLTISFPKPLVDYFANKPMIL-HQIEALKEVGVTEVVLAINRPEVMI 61

Db 22 ALLIAGGRTRLKDLTKRAKPAVHFGGKFRIIDFALSNCINSIRRMGVITQYQSHTLV 81
QY 62 NFLK-----DFEDKLGITITCSQETPLGTAGPLALARDKLDAGSGQFFVILNS 110
Db 82 QHIQRGWSFFNEEMNEFVLLPAQQRKMGENVYRGTAQVNTQNLDIRRYKAEYVVLG 141
QY 111 DIVISEPFAELIKFKHCHGEATI--MVTKVDEPSKYGVVVMEEATGRVERFVEKPKIFV 168
Db 142 DHIYKODYSRMLIDHVEKGVRCVVCMPVPIEASAFGMVAVDENDKTIIE-FVEKP----- 196
QY 169 GNK-----INAGIYLLNPSVL-----DRIELRPTSIEKEVPFQIAADQOQLYA 210
Db 197 ANPPMPNDPSKSLASMGIVYFDADYLYELLEDDRRDENSNDHDFGKDLIPKITEAGLAYA 256
QY 211 MVLP-----GFMDVQPRDYITG-----LRLYDSIRKK 240
Db 257 HPFPLSCVQSDPDAPYWRDVGTLAYWKANLDLASVVPPELMDYRNPRTYNESLPPA 316
QY 241 SAAKATGAHVGNVNLVHESAKIGEGCLIGPDVAIGPCVVEDGVRLSCTVMRGVRIKK 300
Db 317 KFDQDRSGH--GMTL---NSLVSDGCVIS-----GSVVQSVLFSR-----VRVNS 358
QY 301 HACISNSIIGHSTVGQWARIENMTI-----LGEDVHVCDEVY--SNGGVVLPH 347
Db 359 FCNIDSALLPEVWVGRSCLRRRCVDRACVPIEPMGVIGENAEEDARRFYRSEEGIVLVT 418
QY 348 KEI 350
Db 419 REM 421

RESULT 13
US-08-406-858-4
; Sequence 4, Application US/08406858
; Patent No. 5648249
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Stark, David M.
; APPLICANT: Zalewski, James C.
; TITLE OF INVENTION: Method of Improving the Quality of
; TITLE OF INVENTION: Stored Potatoes
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Grace L. Bonner, Monsanto Company, BB4F
; STREET: 700 Chesterfield Parkway No. 5648249th
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/406,858
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/05275
; FILING DATE: 18-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/070,155
; FILING DATE: 28-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Bonner, Grace L.
; REGISTRATION NUMBER: 32,963
; REFERENCE/DOCKET NUMBER: 38-21(10654)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314) 537-7286
; TELEFAX: (314) 537-6047

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 431 AMINO ACIDS
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-406-858-4

Query Match 10.2%; Score 191; DB 1; Length 431;
Best Local Similarity 22.5%; Pred. No. 1.3e-12;
Matches 95; Conservative 60; Mismatches 170; Indels 96; Gaps 15;

QY 3 ALLVCGGCTRLRLPTLSFPKPLVDFAKPMIL-HOIEALKEVGVTEVVLAINVREVM 61
DB 22 AULLAGGRTLRKDLTKRAKPAVHFGGKFRILDFALSNCINSIRRMGVITQYSHTLV 81
QY 62 NPLK-----DFEDKLGITITCSQTEPLGTAGPLALARDKLDGSGQPFVVLNS 110
DB 82 OHIORCWSFFNEEMNEFVLLPAOORMKGENWYRGTAQVTONLDIIRRYKAEYVVLAC 141
QY 111 DVISEYFAELIKFKHKGGEATI--MVTKVDEPSKYGVVMEATGRVERFVEKPKIFV 168
DB 142 DHYKQDYGRMLDHYVEKGVCTVVCMPVIEBASAFGVMAVDENDKTIE-FVEKP---- 196
QY 169 GNK-----INAGIYLLNPSVL-----DRIELRPTSIKEVFPQIAADQOLYA 210
DB 197 ANPPSPNDPSKSLASMGIVFDADLYELLEEDDRDENSNDHDFGKDLIPKITEAGLAYA 236
QY 211 MVLP-----GFMDVQPRDYITG-----LRLYDLSIRKK 240
DB 257 HPFPLSCVSDDAEPYWRDVTGLAYWKANLDLASVPELMDYDRNWPRTVNESLPPA 316
QY 241 SAAKLTAGHYVGNVLYHESAKIGEGCLIGPDVAIGPGCVVDEGVRLSCTVMRGVRIKK 300
DB 317 KFDVRSQSH--GMFL--NSLVSDGCVIS-----GSVVQSVLFSR-----VRVNS 358
QY 301 HACISNIIIGWHSITVCGWARIENMTI-----LGEDVHVCDVY--SNGGVVLPH 347
DB 359 FCNIDSALLPEVWYVGRSCLRCVIDRACVIEGVMVIGENAEEDARRFYRSEEGIVLVT 418
QY 348 KEI 350
DB 419 REM 421

RESULT 14

PCT-US91-04036-4

Sequence 4, Application PC/TUS9104036
GENERAL INFORMATION:
APPLICANT: Kishore, Ganesh M.
TITLE OF INVENTION: Increased Starch Content In Plants
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Monsanto Co.
STREET: 700 Chesterfield Village Parkway
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63198
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/04036
FILING DATE: 19910607
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: McBride, Thomas P.
REGISTRATION NUMBER: 32706
REFERENCE/POCKET NUMBER: 38-21(10530)A
TELECOMMUNICATION INFORMATION:

TELEPHONE: (314) 537-7357
TELEFAX: (314) 537-6047
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 431 AMINO ACIDS
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US91-04036-4

Query Match 10.2%; Score 191; DB 5; Length 431;
Best Local Similarity 22.5%; Pred. No. 1.3e-12;
Matches 95; Conservative 60; Mismatches 170; Indels 98; Gaps 15;

QY 3 ALLVCGGCTRLRLPTLSFPKPLVDFAKPMIL-HOIEALKEVGVTEVVLAINVREVM 61
DB 22 AULLAGGRTLRKDLTKRAKPAVHFGGKFRILDFALSNCINSIRRMGVITQYSHTLV 81
QY 62 NPLK-----DFEDKLGITITCSQTEPLGTAGPLALARDKLDGSGQPFVVLNS 110
DB 82 OHIORCWSFFNEEMNEFVLLPAOORMKGENWYRGTAQVTONLDIIRRYKAEYVVLAC 141
QY 111 DVISEYFAELIKFKHKGGEATI--MVTKVDEPSKYGVVMEATGRVERFVEKPKIFV 168
DB 142 DHYKQDYGRMLDHYVEKGVCTVVCMPVIEBASAFGVMAVDENDKTIE-FVEKP---- 196
QY 169 GNK-----INAGIYLLNPSVL-----DRIELRPTSIKEVFPQIAADQOLYA 210
DB 197 ANPPSPNDPSKSLASMGIVFDADLYELLEEDDRDENSNDHDFGKDLIPKITEAGLAYA 256
QY 211 MVLP-----GFMDVQPRDYITG-----LRLYDLSIRKK 240
DB 257 HPFPLSCVSDDAEPYWRDVTGLAYWKANLDLASVPELMDYDRNWPRTVNESLPPA 316
QY 241 SAAKLTAGHYVGNVLYHESAKIGEGCLIGPDVAIGPGCVVDEGVRLSCTVMRGVRIKK 300
DB 317 KFDVRSQSH--GMFL--NSLVSDGCVIS-----GSVVQSVLFSR-----VRVNS 358
QY 301 HACISNIIIGWHSITVCGWARIENMTI-----LGEDVHVCDVY--SNGGVVLPH 347
DB 359 FCNIDSALLPEVWYVGRSCLRCVIDRACVIEGVMVIGENAEEDARRFYRSEEGIVLVT 418
QY 348 KEI 350
DB 419 REM 421

RESULT 15

PCT-US94-05275-4

Sequence 4, Application PC/TUS9405275
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Method of Improving the Quality of Stored
NUMBER OF SEQUENCES: 26
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/05275
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/070155
FILING DATE: 28-MAY-1993
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 431 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

PCT-US94-05275-4

```
Query Match      10.2%; Score 191; DB 5; Length 431;
Best Local Similarity 22.5%; Pred. No. 1.3e-12;
Matches 95; Conservative 60; Mismatches 170; Indels 98; Gaps 15;

QY 3 ALILVGGFTRRLPLTISFPKPLVDFANKPMIL-HQTEALKEVGTVTEVLAINYRPEVMI 61
    ||||| ||||| || || || || || || || || || || || || || || || || ||
Db 22 ALILAGGRGTRLDLTNKRKPAVHFHGGKFRIIDFALSNCINSIGIRMGVITQYOSHTLV 81
    :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY 62 NFLK-----DFEDKLGITITCSQETEPGLTAGPLALARDKLDGSGQFFVILNS 110
    :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 82 QHIQRGWSFNEEMNEFVDLLPQAQRMKGENWYRGTAQVNTQNLDIIRRYKAETVVILAG 141
    | | : : : : | | : : : : | | : : : : | | : : : : | | : : : :
QY 111 DVISEYPPFAELIKFHKCHGGEATI--MVTKVDEPSKYGVVVMEEATGRVERFVEKPIFV 168
    | | : : : : | | : : : : | | : : : : | | : : : : | | : : : :
Db 142 DHIYKQDYSRMLIDHVEKGYRCHVCMVPVIEASAFGVMAVDENDKTIE-FVEKP---- 196
    :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 169 GNK-----INAGIYLLNPSVL-----DRIELRPTSIEKEVFPQIADQOLYA 210
    : ||| : | | | | | | | | | | | | | | | | | | | | | | | |
Db 197 ANPPSPMPNDFSKSLASNGIYVFDADYLYELLEEDDRDENSDDFGKDLIPKITEAGLAYA 256
    :| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 211 MVL P-----GFWMVVGQPRDYITG-----LRLYLDSTIRKK 240
    :| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 257 HPFPLSCVQSDPDPAEPYWRDVGTLAYWKANLDLASVVPPELDMYDRNWPRTYNESLPPA 316
    :| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 241 SAAKLATGAHVGVNVLVHESAKIGEGCLIGPDVAIGPCYVEDGCVRLSRCVTVMRGVRIKK 300
    :| | : | | | | | | | | | | | | | | | | | | | | | | | | |
Db 317 KFYQDRSGSH--GNTL---NSLVSDGCVIS-----GSVVQSVLFSR-----VRVNS 358
    | :| : | | : | | : | | : | | : | | : | | : | | : | | : | | :
QY 301 HACISNSIIGWHSTVGQWARIENNTI-----LGEDVHVCDDEVY--SNGGVVLPH 347
    | :| : | | : | | : | | : | | : | | : | | : | | : | | : | | :
Db 359 FCNIDSAVLLPEYVWGVGSRCLRCVIDRACVIEGHWIGENAEEDARRFYRSEEGIVLVT 418
    :| : | | : | | : | | : | | : | | : | | : | | : | | : | | :
QY 348 KEI 350
    :| :
Db 419 REM 421
```

Search completed: November 27, 2002, 05:11:25
Job time : 28 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 27, 2002, 05:09:42 ; Search time 38 Seconds
(without alignments)
151.280 Million cell updates/sec

Title: US-09-374-967-2

Perfect score: 1864

Sequence: 1 MKALLVGGFGTRLRLTLTSL.....GVLPHPKEIKSSILKPEIVM 361

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 102317 seqs, 15924203 residues

Total number of hits satisfying chosen parameters: 102317

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	488	26.2	359	9	US-09-712-363-269
2	277.5	14.9	433	10	US-09-925-637-18
3	275.5	14.8	439	10	US-09-815-242-5479
4	275.5	14.8	452	10	US-09-815-242-12149
5	239	12.8	146	10	US-09-738-973-190
6	227	12.2	456	10	US-09-815-242-10400
7	216.5	11.6	458	10	US-09-815-242-11759
8	216	11.6	456	10	US-09-815-242-14064
9	209	11.2	456	10	US-09-815-242-11073
10	190	10.2	292	9	US-09-860-846-12
11	190	10.2	292	10	US-09-861-289-12
12	190	10.2	3782	9	US-09-860-846-4
13	190	10.2	3782	10	US-09-861-289-4
14	189.5	10.2	461	10	US-09-815-242-10937
15	185	9.9	472	10	US-09-815-242-13542
16	179	9.6	479	10	US-09-815-242-13307
17	179	9.6	454	10	US-09-815-242-12112
18	174	9.3	440	10	US-09-734-569-178
19	162.5	8.7	293	10	US-09-934-868-22

Sequence 22, Appl
Sequence 11355, A
Sequence 274, App
Sequence 189, App
Sequence 136, App
Sequence 54, Appl
Sequence 52, Appl
Sequence 53, Appl
Sequence 55, Appl
Sequence 12467, A
Sequence 5707, Ap
Sequence 12467, A
Sequence 248, App
Sequence 10363, A
Sequence 11575, A
Sequence 2, Appli
Sequence 506, App
Sequence 2, Appli
Sequence 1299, Ap
Sequence 13965, A
Sequence 38, Appl
Sequence 576, App
Sequence 39, Appl
Sequence 52, Appl
Sequence 4, Appli
Sequence 144, App
Sequence 449, App

20 162.5 8.7 293 10 US-09-934-868-22
21 142 7.6 433 10 US-09-815-242-11355
22 142 7.6 445 10 US-09-881-752A-274
23 135.5 7.3 160 10 US-09-738-973-189
24 120 6.4 237 10 US-09-765-272-136
25 94 5.0 398 10 US-09-896-852-54
26 94 5.0 506 10 US-09-896-852-52
27 94 5.0 551 10 US-09-896-852-53
28 94 5.0 667 10 US-09-896-852-55
29 93.5 5.0 236 10 US-09-815-242-5707
30 93.5 5.0 242 10 US-09-815-242-12467
31 92 4.9 704 10 US-09-912-020-248
32 92 4.9 704 10 US-09-815-242-10363
33 91 4.9 171 10 US-09-815-242-11575
34 90 4.8 238 10 US-09-876-182-2
35 89.5 4.8 360 10 US-09-841-132-506
36 88.5 4.7 388 10 US-09-938-540-2
37 88 4.7 449 10 US-09-925-301-1299
38 87 4.7 704 10 US-09-815-242-13965
39 86.5 4.6 1981 9 US-09-928-457-38
40 85.5 4.6 354 10 US-09-841-132-576
41 84.5 4.5 608 10 US-09-767-041-39
42 84 4.5 57 10 US-09-734-017A-52
43 84 4.5 182 10 US-09-734-569-4
44 84 4.5 184 10 US-09-734-569-144
45 84 4.5 1252 10 US-09-841-132-449

ALIGNMENTS

RESULT 1

US-09-712-363-269

; Sequence 269, Application US/09712363

; Patent No. US20020164588A1

; GENERAL INFORMATION:

; APPLICANT: Eisenberg, David

; APPLICANT: Rotstein, Sergio H.

; APPLICANT: Marcotte, Edward M.

; TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND

; FILE OF INVENTION: INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS

; FILE REFERENCE: 07419-032001

; CURRENT APPLICATION NUMBER: US/09/712,363

; CURRENT FILING DATE: 2000-11-13

; PRIOR APPLICATION NUMBER: PCT/US00/02246

; PRIOR FILING DATE: 2000-01-28

; PRIOR APPLICATION NUMBER: 60/179,531

; PRIOR FILING DATE: 2000-02-01

; PRIOR APPLICATION NUMBER: 60/117,844

; PRIOR FILING DATE: 1999-01-29

; PRIOR APPLICATION NUMBER: 60/118,206,

; PRIOR FILING DATE: 1999-02-01

; PRIOR APPLICATION NUMBER: 60/126,593

; PRIOR FILING DATE: 1999-03-26

; PRIOR APPLICATION NUMBER: 60/134,093

; PRIOR FILING DATE: 1999-05-14

; PRIOR APPLICATION NUMBER: 60/134,092

; PRIOR FILING DATE: 1999-05-14

; PRIOR APPLICATION NUMBER: 60/165,124

; PRIOR FILING DATE: 1999-11-12

; PRIOR APPLICATION NUMBER: 60/165,086

; PRIOR FILING DATE: 1999-11-12

; NUMBER OF SEQ ID NOS: 292

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 269

; LENGTH: 359

; TYPE: PRT

; ORGANISM: Mycobacterium tuberculosis

; US-09-712-363-269

Query Match

Best Local Similarity 26.2%; Score 488; DB 9; Length 359;

Matches 125; Conservative 63; Mismatches 151; Indels 22; Gaps 10;

```

QY 3 ALLIVGGFGTGLRPLTLSPKPLVDFAKPMILHOIEALKEVCVTEVVLAINVRPEVMIN 62
DB 8 AVVLVGGGKGTGLRPLTLSPKPLVDFAKPMILHOIEALKEVCVTEVVLAINVRPEVMIN 64
QY 63 FLKDFED--KLGIITWCSETEPLCPAGLALARDKLDGSCQPFVLSNDVSEVFAE 120
DB 65 FRAEFGDSALGLQIEYVTEHPGLTGGGIANVAGKLNRDTA---WVFGDVLUSGADLAQ 121
QY 121 LKFKHCHGGEATIMVTKVDEPSKYGVVYVMEATGRVERFVKFIYFGNKNINAGIYLLN 180
DB 122 LLDFHRSNRADVTLOLVRCDDPRAFCVPTDE-EDRVVAFLEKTEDDPDQINAGCVVPE 180
QY 161 PVLVDRI--ELRSTSEKEVFPQIADQ--OLYAMVLPFGWMDVGPQPRDITIGLRLYLDISI 237
DB 181 RNVIDRIPOGREVSVERVEFPALLADGCKIKYGVDAWYRDMGTPEDEFVRG---SADLV 237
QY 238 RKKSAAKLATGAHVGNVLAVHESAKIGECCLIGCPDVAIGPCCVVEDGVLRSCTVMRGVR 297
DB 238 RGIAFSPALRGHR--GEQJYHDAAYSPGALLIGGTVVGGAIEGPTRLDGGAVIFDGV 295
QY 208 IKKHACISNIIGWSTVGMARIENNTI--LGEDVHVCDVYSNG---GVVLPKHEIKS 352
DB 296 VEAGCVIERSIIGFAGRICPRALIRDVIGDGDADICARCELLSGARWPGVFLPDGGIRY 355
QY 353 S 353
DB 356 S 356

RESULT 2
us-09-925-637-18
; Sequence 18, Application us/0925637
; Patent No. US2002010338A1
; GENERAL INFORMATION:
; APPLICANT: CHOI
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Polypeptides
; FILE REFERENCE: PB560
; CURRENT APPLICATION NUMBER: US/09/925,637
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/23773
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 60/151,933
; PRIOR FILING DATE: 1999-09-01
; PRIOR APPLICATION NUMBER: US 08/781,986
; PRIOR FILING DATE: 1997-01-03
; PRIOR APPLICATION NUMBER: US 08/956,171
; PRIOR FILING DATE: 1997-10-20
; PRIOR APPLICATION NUMBER: US 60/009,861
; PRIOR FILING DATE: 1996-01-06
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Homo sapiens
us-09-925-637-18

```

Query Match 14.9%; Score 277.5; DB 10; Length 453;
 Best Local Similarity 27.2%; Pred. NO. 1.2e-19;
 Matches 106; Conservative 64; Mismatches 146; Indels 73; Gaps 15;

```

QY 3 ALLIVGGFGTGLRPLTLSPKPLVDFAKPMILHOIEALKEVCVTEVVLAINVRPEVMIN 62
DB 8 AVVLVGGGKGTGLRPLTLSPKPLVDFAKPMILHOIEALKEVCVTEVVLAINVRPEVMIN 64
QY 63 FLKDFEDKLGIITTCSETEPLCTAGPLALARDKLDGSCQPFVLSNDVSEVFAEL 121
DB 61 ---SVKHLGERSLSYFQDEQLGTAHAYOMAKSHLEDKEGTTIVVCGDTPLIITKELTVL 117
QY 122 IKFKHCHGGEATIMVTKVDEPSKYGVVYVMEATGRVERFVKFIYFGNKNINAGIY 177
DB 118 IAHEDANAQATVLSASIQQPYGGRIV--RNASGRLEIVEEKDAQAEKDINEISSGIF 176

```

```

QY 178 LLNPSVLDRIELRPTSTSEKEVFPQIADQIYAMVLP-----GFWMVDVGQPRDY-- 226
DB 177 AFNNKTL-----FEKLTQVKNDNAOGEYYLDPVLISLINDGGIVEYRTNDVVEE 225
QY 227 ITGL--RLYLDSTIRKSAAKLATGAHVYGNVLYHESAKIGEGCLIGPYALGPGCVYVEDG 284
DB 226 IMCVNDRVML--SQEKAMORRTNHYHMLNGVTIIDP-----DSTVIGDDVTICSDVTIEPG 280
QY 285 VRLS-----RCTVMRGYRIKKHACISNSI-----IGHWSTVGMARIENMTIL 327
DB 281 VRINGRTEIGEDVVYIGQYSEINNSTIENGACIQGVVNDASVGANTKVGPFPQALRPGAQL 340
QY 328 GEDVHVCDVYSNCGVVLPHKEIKSSILK 356
DB 341 GADYKV-----GNFY-----EIKKADLK 358

RESULT 3
US-09-815-242-5479
; Sequence 5479, Application us/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 5479
; LENGTH: 449
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5479

```

Query Match 14.8%; Score 275.5; DB 10; Length 449;
 Best Local Similarity 27.2%; Pred. NO. 1.9e-19;
 Matches 106; Conservative 64; Mismatches 146; Indels 73; Gaps 15;

```

QY 3 ALLIVGGFGTGLRPLTLSPKPLVDFAKPMILHOIEALKEVCVTEVVLAINVRPEVMIN 62
DB 4 AIIAAGKGTMRKSKY---KYLHEVAGKPMVHELVESKVGSGVDQVVTIVHGAE----- 56
QY 63 FLKDFEKLGIITTCSETEPLCTAGPLALARDKLDGSCQPFVLSNDVSEVFAEL 121
DB 57 ---SVKHLGERSLYSPDEQLGTAHAYOMAKSHLEDKEGTTIVVCGDTPLIITKELTVL 113
QY 122 IKFKHCHGGEATIMVTKVDEPSKYGVVYVMEATGRVERFVKFIYFGNKNINAGIY 177
DB 114 IAHEDANAQATVLSASIQQPYGGRIV--RNASGRLEIVEEKDAQAEKDINEISSGIF 172

```



```
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10400
; LENGTH: 456
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-815-242-10400
```

```
Query Match 12.2%; Score 227; DB 10; Length 456;
Best Local Similarity 24.4%; Pred. No. 13e-14;
Matches 100; Conservative 55; Mismatches 158; Indels 96; Gaps 16;

QY 1 MKALILVGFGTRRLRLTLSPKPLVDFAKPKMLHQIEALKEVGYEVTVVLAIRPEVM 60
Db 6 MSVVILAAGKGRMYS---DLPRVLTLAGKAMVQHVDAANELAAHVHLVYGHGGDLL 62

QY 61 INFLKDFEDKLGITITCSQETEPGLTAGPLALARDKLADGSGOPFFVLNSDVISEYPPAE 120
Db 63 KOALKD--DNLNWL---QAEQLGTGHAMQQA-----AFFTADDEDILMLYGDVP 107

QY 121 LKFKH-----KCHGEATIMTKVDERSKYGVVYME--EATGRVER--FVEKPKIF 167
Db 108 LLSVETLQRLRAKPOGG--IGLLTVKLDPTGYGRITRENGKVTGIVEHKDAIDEQOI- 165

QY 168 YGNKINAGIYLLN-----PSVLDRI 187
Db 166 --QEINTGILIANGADMKRWLAKLTNNNAQGEVYITDIALAVQEGREIVAVHPQLSEV 223

QY 188 ELRPTSIEKEVPQTAADQOLYAMVLPGFWMVGQPRDYIT-GLRVLDSIRKKSAAKLA 246
Db 224 EGVNRLQLSLRLERYVQSEQAELKLLAGVNL-----RDFAREDLRGTLTHGRD---VEID 275

QY 247 TGAHVGVNVLVHESAKIGEC-----LIGPDVAIGCCVVEDGVRLSRCTV-----MRC 295
Db 276 TFWIIEGNTGLHRVKIGTCVTKNSVIGDCEISPTVVEDANLAAACTIGPFARLRP 335

QY 296 VRIKKHACISNLSIGHVSTVGOWARTENMTILGEDVHVCDEVYSGNGVV 344
Db 336 ALLEGAHVGNFVEMKKARLGKSKAGHLYLG--DAEIGNVNIAGTIC 383
```

```
RESULT 7
US-09-815-242-11759
; Sequence 11759, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
```

```
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11759
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-815-242-11759
```

```
Query Match 11.6%; Score 216.5; DB 10; Length 458;
Best Local Similarity 23.6%; Pred. No. 1.5e-13;
Matches 100; Conservative 62; Mismatches 156; Indels 105; Gaps 17;

QY 1 MKALILVGFGTRRLRLTLSPKPLVDFAKPKMLHQTALKEVGYEVTVVLAIRPEVM 60
Db 8 MSVVILAAGKGRMYS---DLPRVLTLAGKPMYQHVDAANDLACAVHLVYGHGGDLL 64

QY 61 INFLKDFEDKLGITITCSQETEPGLTAGPLALARDKLADGSGOPFFVLNSDVISEYPPAE 120
Db 65 RQTL--HEDNLNWL---QAEQLGTGHAMQQA-----APFFNDEDEILMLYGDVP 109

QY 121 LKFKH-----KCHGEATIMTKVDERSKYGVVYME--EATGRVEREVEKPKIFVN 170
Db 110 LISVETLQRLRAKPOGG--IGLLTVKLDPTGYGRITRENGKVTGIVEHKDAISEAQROI 168

QY 171 KINAGIYLLN-----LNFSVLDRIE-- 188
Db 169 EINTGILIANGADLKRWLAKLTNNNAQGEYITDIAMAHQEGHIVAVHPQLSEVEGV 228

QY 189 ---LRPTSIEKEVPQTAADQOLYAMVLPGFWMVGQPRDYITGLRLVLDIRKKSAAKL 245
Db 229 NNRLQALRLER-VYQAEQAEKLLLAGVNL-----LRDPARF--DLR---GLTQHGROVEI 276

QY 246 ATGAHVGVNVLVHESAKIGEC-----IGPDVAIGCCVVEDGVRLSRCTV-----MR 294
Db 277 DTNVILGNGVGLDRVKIGACGVTKNSTIGDDCEISPTVSVVEDAQLHAACITIGPFARLRP 336

QY 295 GVRIKKHACISNLSIGHVSTVGOWARTENMTILGEDVHVCDEVYSGNGVV-----PH 347
Db 337 GAELLEGAHVGNFVEMKKARLGKSKAGHLYLG--DAEIGNVNIAGTICVNDCAANKH 395

QY 348 KEI 350
Db 396 KTI 398
```

```
RESULT 8
US-09-815-242-14064
; Sequence 14064, Application US/09815242
; Patent No. US20020061569A1
```

```

RESULT 9
US-09-815-242-11073
; Sequence 11073, Application US/09815242
; Patent NO. US2002061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11073
; LENGTH: 456
; TYPE: prt
; ORGANISM: Haemophilus influenzae
US-09-815-242-11073

```

```

Query Match      11.2%; Score 209; DB 10; Length 456;
Best Local Similarity 23.5%; pred. No. 8.1e-13;
Matches 95; Conservative 63; Mismatches 159; Indels 88; Gaps 15;

QY 1 MKALILVGGFGRRLPTLSFPKPLVDFAKXPMILHQIEALKEGCVTEWVLAINRPEVM 60
   : : : | | | | | : | | | : | : : | : : : | : : |
Db 6 LSAVILAAGKGRTRYMS---DLPKVLHTTAGKPMWKHVDTAHQLGSENIHLITYHGDDLM 62

QY 61 INFLKDFEDKLGITITCSQETPLCTAGPLALARDKLADGSGQPFVLSNSOISISYPFAE 120
   : : : | | | | | : | | | | : : : | : : : |
Db 63 RTHLANEQ-----VNMVLQTEQLGTGAHAVQQA-----APFKDNEINVLVYGDAP 107

QY 121 LIKPH-----KCHGGEATIMTKVDEPSKYGVVMEEATGRVERVE-----162
   | | | | | | | | | | | | | | | | | | | | | | |
Db 108 LITKETLEKLEAKPENGLA--LLAVNLNDNPTGYGRIIRE--NGNVVAIVEQDANAEOQLN 164

QY 163 -----KPKIFVGNKINAGIYLLNP--SVLDRIELRPTSTIE-KEVFPQI 202
   | | | | | | | | | | | | | | | | | | | | |
Db 165 IKEVNTGMVNSDGFASKFKWLARVGNNAQGEYLLDILLANQDNCQVVAVQATDVMVE 224

QY 203 AADQOLYAMVLPGFWMQVGPQDXYITGLURLYLDSTIRKKSAAKLATGAHV-----VGN 254
   | : | : | : | : | : | : | : | : | : | : | : |
Db 225 GANNRLQALALERYFQNKQASKLLEGVMYI-DPARFDLRGTLEHGKDVEIDVNVYIEGN 283

QY 255 VLVHESAKIGBC-----LIGPDVAIGPGCVVEGDGVRLSRCTV-----MRGVRIKKHAC 303
   | : | : | | | | | : | | | | | | | | | | |
Db 284 VKLGDVRKVIQGCVLKNVIVINDVEIKRYPVLEDSIVGEKAAIGPFSRLRPGCAELAEETH 343

QY 304 ISNSIIHGHWSTVGQWARIENNTILGEDVHVHDEYVSN-----GGVV 344
   | | | | | | | | | | | | | | | | | | | | |
Db 344 VGNFEVEIKTSIVGKSGKNVHITYYGD-----SEIGSNCNIGAGVI 383

```

[illegible]

```

; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10937
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-10937

Query Match          10.2%   Score 189.5;   DB 10;   Length 461;
Best Local Similarity 21.5%;   Pred. No. 7.1e-11;
Matches 87;   Conservative 70;   Mismatches 158;   Indels 89;   Gaps

Qy 3  ALILVGGFGRRLRPLTISFPKPLVDFAFKPMILHQIBALKEVGVTEVVLAINRYRPEVMIN 62
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9  AIIAAGKGRMKS--KLYKVLHPVAGKPMVEHILDQVEQTEPTIETIVGHGAEMI-- 63

Qy 63 FLKDFEDKLGITITCSQFTEPLGTAGLALARDKLADGSGQPPFVL-NSDVISYPPFAEL 121
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 64 -----KSHLGRSQAALQAEOLGTGHAVMQAQLLGGKQGTLLVITGDTPLLTAAETLKNL 118

Qy 122 IKPHKCHGGEATIMVTKVDEPSKGVVVMEEATGRVERFEKPKIFVG-----NKINAGIY 177
      :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 119 FDYHQGNASATILTAHAEDPTGGRRIIRDH-VGIVERIVEQKDASEEARVOEINTGTF 177
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 178 LL-NPSVLDRIELRPTS-----IEKVFQIADQOQLYAMVLPGFWMMDVG-- 221
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 178 CFDNESLFEALAKTDTNNTQGEYVLTDIIEILKEGKAVAAQY-----MADFEAMGVN 231

Qy 222 -----QPRDYITGL---RLYDSIRKKSAAKLATGAHVGVNVLVHE 259
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 232 DRVALSTANKIMHRRLENMHRNGVTFIDPDTYID-----EGVVGSDTVIEAGVTIKG 286

Qy 260 SAKIGECCLIGP-----DVAIGPCVVVEDGVRLSRCTVMRGVRIKKHACI-SNSIIGWHS 313
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 287 KTVIGEDCLIGAHSEIVDSHIGNVVVQKSV-IEESVVRGADVGPYAHLRPRADVGANV 345

Qy 314 TVGQWARIENNTI-----LGEDVHV-CDEVYSN 340
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 346 HIGNFVEVKNATIDEGTKVGHLYTVGVDATLGKDINVCGVVFEVN 389

RESULT 15
US-09-815-242-13542
; Sequence 13542, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haseibeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA 011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625

```

```
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/369,308
/ PRIOR FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 14110
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 13542
/ LENGTH: 472
/ TYPE: prt
/ ORGANISM: Streptococcus pneumoniae
US-09-815-242-13542

Query Match      9.9%; Score 185; DB 10; Length 472;
Best Local Similarity 24.3%; Pred. NO. 2.1e-10;
Matches 90; Conservative 56; Mismatches 175; Indels 50; Gaps 11;

QY 3 ALLILGCGFTRLRPLILSPFKPLVDFAFKPMILHQIEALKEVGYTEVYVLAINRPEVMIN 62
DB 18 AIIAAGKGRKMS---DLFKVLHKVAGISMLEHVFERSGAIQPEKTVTVVGHKAELV-- 72
QY 63 FLKDPEDKLGITITCSQETPEPLCTAGPLALARDKLADSGOPFFVL-NSDVISEYPPFAEL 121
DB 73 -----EEVLAEQTEFTQSEQLGIGHVMMTEFLEGLSGHTLVIAAGDTPLITGESLKNL 127
QY 122 IRFKHCHGGEATIMVTKVDPSKYGVVVMEEATCRVERFVEKPKIFVGNK-----INAGIY 177
DB 128 IDFHINHKNVATILTAETDNPFYGRIVRND-NAEVLRIVEQKDATDFEKQKEINTGY 186
QY 178 LL-NPSYLDRIELRPTSIEKEVFFQIAADQQLYAMVLPGFMDVGCQPRDYITGLRLYLD 236
DB 187 VFDNERLFEALKNINFN-----NAOGEYVITDVIGIFRETGEKVGAYT-LKDFDES 236
QY 237 IRKKSAAKLATGAHVGNVLVHESAKIG-----EGCLIGPDVAIGPGCVVEDGV----- 285
DB 237 LGVNDRVALATAESVMRRRINHRHMVNGSVFVPEATYIDIDVETIADPEVOIEANVILKQ 296
QY 286 -----RLSRCTVMRGVRIKKHACISNSIICWHSS-----TVGOWARIENMTILGEDVHV 333
DB 297 TKIGAETVLTNGTYVVDSTIGAGAVITNSMIESSVADGVTVGPIAHIRENSSLGQVHI 356
QY 334 CDEVSYNGGVV 344
DB 357 GNFVEVKGSSI 367
```

Search completed: November 27, 2002, 05:19:00
JOB time : 42 secs

Db 181 PSVLKDIELRPSIEKETPPKIAAAGCLYAMVLPGFWMIDGQPRDITGLRLYDLSURKK 240
Qy 241 SAAKLATGAHVYGVNVLVHESAKIGECGLICPDVAIGPGCVVEDGVRLSRTVMRGVRIKK 300
Db 241 SPAKLTSGPHIYGVNLYDETATIGEGCLIGPDVAIGPGCVESGVRLSRTVMRGVRIKK 300
Qy 301 HACISNSIIGWSTVCGWARIENMTILGEDVHVHCVDEYNGVGVLPKHKEIKSSILKPEIV 360
Db 301 HACISSIIIGWSTVCGWARIENMTILGEDVHVHSDRIYSNGGVLPKHKEIKSNILKPEIV 360
Qy 361 M 361
Db 361 M 361

RESULT 2
T47698
mannose-1-phosphate guanylyltransferase-like protein - Arabidopsis thaliana
N:Alternate names: protein T22E16.250
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 19-May-2000
C:Accession: T47698
R:Benes, V.; Wurmbech, E.; Drzonek, H.; Ansoorge, W.; Lemcke, K.; Mayer, K.F.
submitted to the Protein Sequence Database, February 2000
A:Reference number: 224472
A:Status: preliminary
A:Map position: 4
A:Molecule type: DNA
A:Residues: 1-364 <BEN>
A:Cross-references: EMBL:AL132975
A:Experimental source: cultivar Columbia; BAC clone T22E16
C:Genetics:
A:Map position: 3
A:Introns: 37/3: 58/3: 141/3
A:Note: T22E16.250
C:Superfamily: mannose-1-phosphate guanylyltransferase

Query Match 85.2%; Score 1567.5; DB 2; Length 364;
Best Local Similarity 83.2%; Pred. No. 2.4e-118;
Matches 303; Conservative 29; Mismatches 29; Indels 3; Gaps 1;

Qy 1 MKALILVGGFGTRLRPLTSLSPFKPLVDFANKPMILHQIEALKEVGYTEVVLAINYPE-- 58
Db 1 MKALILVGGFGTRLRPLTSLPKPLVDFANKPMILHQIEALKAIGDEVVLAINYPEQL 60
Qy 59 -VMINFLKDFEDKIGITITCSQETPLGTAGPLALARDKLADGSGQFFVLSNDVISEYP 117
Db 61 LVMSKFSNDVEATLGKIKITCSQETPLGTAGPLALARDKLVDGSGQFFVLSNDVISEDY 120
Qy 118 PAELIKPHKCHGGEATIMVTKVDEPSKGVVWMEETGVRERVEKPIFVGNKINAGIV 177
Db 121 LEEMTAFNHAAGEASIMVTKVDEPSKGVVWMEETGVRERVEKPIFVGNKINAGIY 180
Qy 178 LNPVSLDRIELRPTSIKEIEFPQIAAEKLYAMLPFGFWMIDGQPRDYITGLRLYDLSI 237
Db 181 LNPVSLDRIELRPTSIKEIEFPQIAAEKLYAMLPFGFWMIDGQPRDYITGLRLYDLS 240
Qy 238 RKSAKLATGAHVYGVNVLVHESAKIGECGLIGPDVAIGPGCVVEDGVRLSRTVMRGVR 297
Db 241 RKSSKPLATGPHILGNVLVDYETAEIGEGCLIGPVAIGPGCVYEGYRLSHCTYHRGVH 300
Qy 298 IKKHACISNSIIGWSTVCGWARIENMTILGEDVHVHCVDEYNGVGVLPKHKEIKSSILKP 357
Db 301 VARYACISSIIIGWSTVCGWARIENMTILGNVYVCDIYCNCGVVLHNKKEIKSDILKP 360
Qy 358 EIVM 361
Db 361 DIVM 364

RESULT 3
F85357

GDP-mannose pyrophosphorylase like protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 02-Mar-2001
C:Accession: F85357
R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sp
Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083489; PMID:10617198
A:Accession: F85357
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-351 <STO>
A:Cross-references: GB:NC_001268; NID:g7269958; PIDN:CAB79775.1; GSPDB:GN00140
C:Genetics:
A:Gene: AT4g30570
A:Map position: 4
C:Superfamily: mannose-1-phosphate guanylyltransferase

Query Match 73.8%; Score 1376; DB 2; Length 351;
Best Local Similarity 76.3%; Pred. No. 1.4e-101;
Matches 264; Conservative 44; Mismatches 36; Indels 2; Gaps 2;

Qy 1 MKALILVGGFGTRLRPLTSLSPFKPLVDFANKPMILHQIEALKEVGYTEVVLAINYPEEV 59
Db 1 LKALILVGGFGTRLRPLTSLSPFKPLVDFANKPMILHQIEALKEVGYTEVVLAINHQOPEV 60
Qy 60 MINFLKDFEDKIGITITCSQETPLGTAGPLALARDKLADGSGQFFVLSNDVISEYPFA 119
Db 61 MLNPFVKEYEKLKIKITFSQETPLGTAGPLALARDKLVDGSGQFFVLSNDVISEYPLL 120
Qy 120 ELIKFKHCHGGEATIMVTKVDEPSKGVVWMEETGVRERVEKPIFVGNKINAGIYLL 179
Db 121 EMIEFHKNRAEASIMVTEVDDPSKYGVVTEGTARVESFEVKPKHFVGNKINAGIYLL 180
Qy 180 NPSVLDRIELRPTSIKEIEFPQIAADQOLYAMVLPFGFWMIDGQPRDYITGLRLYDLSIRK 239
Db 181 SPVSLDRIELRPTSIKEIEFPQIAAEKLYAMVLPFGFWMIDGQPRDYITGLRLYDLSLRE 240
Qy 240 KSAKLATGAHVYGVNVLVHESAKIGECGLIGPDVAIGPGCVVEDGVRLSRTVMRGVRIK 299
Db 241 KTOELATGNIIGNVLVHESAVICEGLIGPDVVIIGPGCVDSGVRLFCGTVMRGVWIK 300
Qy 300 KHCISNSIIGWSTVCGWARIENMTILGEDVHVCD-EVYSNGGVV 344
Db 301 EHACISNSIVGDSVTGWRARVENITVLGKDVNVADADEVYNSGVYI 346

RESULT 4
S67590
mannose-1-phosphate guanylyltransferase (EC 2.7.7.13) - yeast (Saccharomyces cerevisi
N:Alternate names: protein D2566; protein YDL055c; PSA1 protein
C:Species: Saccharomyces cerevisiae
C:Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 20-Jun-2000
C:Accession: S67590; S59373; S59374; S62744
R:Blocker, H.; Brandt, P.
submitted to the Protein Sequence Database, July 1996
A:Reference number: S67587
A:Accession: S67590
A:Molecule type: DNA
A:Residues: 1-361 <BLO>
A:Cross-references: EMBL:Z74103; NID:g1431052; PIDN:CAA98617.1; PID:g1431053; MIPS:YD
A:Experimental source: strain S288C
R:Schultz, J.; Sprague, G.F.
submitted to the EMBL data Library, April 1995
A:Description: Isolation of the gene encoding mannose-1-phosphate guanylyltransferase f
A:Reference number: S59737
A:Accession: S59737
A:Molecule type: DNA
A:Residues: 1-49, 'A', 51-361 <SCH>
A:Cross-references: EMBL:U24437; NID:g894203; PIDN:AA659677.1; PID:g894204
R:Benton, B.K.; Plump, S.D.; Roos, J.; Lennarz, W.J.; Cross, P.R.
submitted to the EMBL data Library, January 1995
A:Description: Overexpression of 5. Cerevisiae G1 cyclins restores viability of algi

A;Reference number: S59374

A;Accession: S59374

A;Molecule type: DNA

A;Residues: 1-15,'G',17-361 <BEN>

A;Cross-references: EMBL:U19608

R:Benton, B.K.; Plump, S.D.; Roos, J.; Lennarz, W.J.; Cross, F.R.

Curr. Genet. 29, 106-113, 1996

A;Title: Over-expression of *S. cerevisiae* G(1) cyclins restores the viability of *alg1* N-

A;Reference number: S62744; MUID:96418871; PMID:8821656

A;Accession: S62744

A;Status: nucleic acid sequence not shown

A;Molecule type: DNA

A;Residues: 1-361 <BEN>

A;Cross-references: EMBL:U19608; NID:g1292897; PIDN:AAC49289.1; PID:g1292898

C;Genetics:

A;Gene: SGD:PSA1; MPG1

A;Cross-references: SGD:S0002213; MIPS:YDL055c

A;Map position: 4L

C;Superfamily: mannose-1-phosphate guanylyltransferase

C;Keywords: nucleotidyltransferase

Query Match 63.2%; Score 1178; DB 2; Length 361;
Best Local Similarity 61.0%; Pred. No. 7.5e-86;
Matches 221; Conservative 56; Mismatches 83; Indels 2; Gaps 2;

QY 1 MKALILVGGFGRRLRPLTLSPKPLVDFAFKPMILHQIEALKEVGTVTVLAINRPEVM 60

Db 1 MKGLILVGGFGRRLRPLTLVPKPLVDFGNRPMLHQIEALANAGVTDIVLAVNYRPEVM 60

QY 61 INFLKDFEDKLGITTCQETEPGLTAGLALARDKLDGSGQPFVFLNSDVISEYPAE 120

Db 61 VETLKKEKEGVNITVSVEPELGTAGPLAEDVLKK-DNSPFFVFLNSDVISEYPAE 119

QY 121 LIKFKHCHGGEATIMVTKVDEPSKYGVVMEAT-GRVERFEVEKPKIFVGNKINAGIYLL 179

Db 120 LADPHKAGGKGTATVKVDEPSKYGVIVHDIAENPLNLDREVKEPKFVGNRINAGIYLL 179

QY 180 NPSVLDRIELRPTSTIEKEVFPQIAADQOLYAMVLPFGFMDVGQPRDYITGLRLYLDSTRK 239

Db 180 NPEVIDIEMKPTSTIEKETPFLVEEKQLYSFDELEGFMDVGQPKDFLSGVTVLYLSLAK 239

QY 240 KSAAKLATGARVGNVLVHESAKIGEGCLIGPDVAIGPCVVEDGVRLSRTVMRGVRIK 299

Db 240 RQPKLATGANIVGNALIDPTAKTSSTAKIGPDVVGPNVTIGDVRITRSVVLNCSSTIK 299

QY 300 KHACISNIIGWSTVGOWARIENMTILGEDVHVCDEYVNSGVVLPKHEIKSSILKPEI 359

Db 300 NLSLVKSTIVGNSTVGQWCRLEGVTVLGDDEVEKDEIYINGGVLPKHSISDNVPEAI 359

QY 360 VM 361

Db 360 IM 361

RESULT 5

T41209

mannose-1-phosphate guanylyltransferase - fission yeast (Schizosaccharomyces pombe)

C;Species: Schizosaccharomyces pombe

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000

C;Accession: T41209

R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Murphy, L.; Harris, D.

submitted to the EMBL Data Library, September 1998

A;Reference number: Z21969

A;Accession: T41209

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-363 <LYN>

A;Cross-references: EMBL:AL031538; PIDN:CAA20770.1; GSPDB:GN00068; SPDB:SPCC1906.01

A;Experimental source: strain 972h-; cosmid c1906

C;Genetics:

A;Gene: SPDB:SPCC1906.01

A;Map position: 3

C;Superfamily: mannose-1-phosphate guanylyltransferase

Query Match 63.0%; Score 1175; DB 2; Length 363;

Best Local Similarity 62.1%; Pred. No. 1.3e-85;

Matches 226; Conservative 54; Mismatches 80; Indels 4; Gaps 4;

QY 1 MKALILVGGFGRRLRPLTLSPKPLVDFAFKPMILHQIEALKEVGTVTVLAINRPEVM 60

Db 1 MKALILVGGFGRRLRPLTLTPKPLVDFGNKPMILHQVEALAAAGVTDIVLAVNYRPEIM 60

QY 61 INFLKDFEDKLGITTCQETEPGLTAGLALARDKLDGSGQPFVFLNSDVISEYPAE 120

Db 61 VEALKKYEKEYNVNITVSVEPELGTAGPLALARDILAK-DHSPFFVFLNSDVISEYPAE 119

QY 121 LIKFKHCHGGEATIMVTKVDEPSKYGVVMEATGRVERFEVEKPKIFVGNKINAGIYLL 179

Db 120 LAAFHKAHGAEGTIVVTKEEPSKYGVVHYPNSESLIERFEVEKPKIFVGNKINAGIYLL 179

QY 180 NPSVLDRIELRPTSTIEKEVFPQIAADQOLYAMVLPFGFMDVGQPRDYITGLRLYLDSTRK 239

Db 180 NPSVLDRIELRPTSTIEKEVFPAMVNDKQLHSPFDELEGFMDVGQPKDYLTGTCLYLSLRK 239

QY 240 KSAAKLATGARVGNVLVHESAKIGEGCLIGPDVAIGPCVVEDGVRLSRTVMRGVRI 298

Db 240 HKPEILAPASSNLIIGNVLIDPSATIGKNCKIGPNVVGPNVTIGDGVRLQRCALIKSSRV 299

QY 299 KHACISNIIGWSTVGOWARIENMTILGEDVHVCDEYVNSGVVLPKHEIKSSILKPEI 357

Db 300 RQHWYKSSIVGNSTVGLSWSLENVSLGDDVYVNDIYVNGSILPHKHSISANIEVPG 359

QY 358 EIVM 361

Db 360 TIVM 363

RESULT 6

T42371

probable mannose-1-phosphate guanylyltransferase (EC 2.7.7.13) - fission yeast (Schiz

C;Species: Schizosaccharomyces pombe

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jul-2000

C;Accession: T42371

R:Yoshioka, S.; Kato, K.; Nakai, K.; Okayama, H.; Nojima, H.

DNA Res. 4, 363-369, 1997

A;Title: Identification of open reading frames in Schizosaccharomyces pombe cDNAs.

A;Reference number: Z17323; MUID:9816272; PMID:9501991

A;Accession: T42371

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-363 <YOS>

A;Cross-references: EMBL:D89128; NID:g1749463; PIDN:BAAL3790.1; PID:g1749464

A;Experimental source: strain PR745

C;Function:

A;Description: generates GDP-mannose and pyrophosphate from mannose-1-phosphate and G

C;Superfamily: mannose-1-phosphate guanylyltransferase

C;Keywords: nucleotidyltransferase

Query Match 62.4%; Score 1164; DB 2; Length 363;

Best Local Similarity 61.5%; Pred. No. 9.7e-85;

Matches 224; Conservative 55; Mismatches 81; Indels 4; Gaps 4;

QY 1 MKALILVGGFGRRLRPLTLSPKPLVDFAFKPMILHQIEALKEVGTVTVLAINRPEVM 60

Db 1 MKALILVGGFGRRLRPLTLTPKPLVDFGNKPMILHQVEALAAAGVTDIVLAVNYRPEIM 60

QY 61 INFLKDFEDKLGITTCQETEPGLTAGLALARDKLDGSGQPFVFLNSDVISEYPAE 120

Db 61 VEALKKYEKEYNVNITVSVEPELGTAGPLALARDILAK-DHSPFFVFLNSDVISEYPAE 119

QY 121 LIKFKHCHGGEATIMVTKVDEPSKYGVVMEATGRVERFEVEKPKIFVGNKINAGIYLL 179

Db 120 LAAFHKAHGAEGTIVVTKEEPSKYGVVHYPNSESLIERFEVEKPKIFVGNKINAGIYLL 179

QY 180 NPSVLDRIELRPTSTIEKEVFPQIAADQOLYAMVLPFGFMDVGQPRDYITGLRLYLDSTRK 239

Db 180 NPSVLDRIELRPTSTIEKEVFPQIAADQOLYAMVLPFGFMDVGQPRDYITGLRLYLDSTRK 239

Db 180 NPSVLDRIBPRPTSIEKEVFPAMVNDKQLHSFDLEGYMDVQPKDYLRGTCTCLYLSLRK 239
QY 240 KSAAKLA-TGAHVGNVLVHESAKIGEGCLIGPDVAIGPCVVEDGVRLSRCTVMRGVRI 298
Db 240 HPEILAPASSNIGNVLDIPATICKKIGENVYVIGNYIGDGYRQRCALIKSSRV 239
QY 299 KKHACISNIGHSTVSGQWARIENMTILGEDVHVDEYVSGGVVLPKHKESSILKP- 357
Db 300 RDHAWKSSIVGNWSTLGSWSRLNVSGLDDVVVNDIYVNGSGSLPHKKSISANIEVPG 359
QY 358 EIVM 361
Db 360 TIVM 363
RESULT 7
T32900
hypothetical protein C42C1.5 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jan-2000
C:Accession: T32900
B:Murray, J.; Rohlfing, M.; Davidson, S.
submitted to the EMBL Data Library, January 1998
A:Description: The sequence of *C. elegans* cosmid C42C1.
A:Reference number: Z1243
A:Accession: T32900
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-373 <KUR>
A:Cross-references: EMBL:AF043695; PIDN:AB97947.1; GSPDB:GN00019; GESP:C42C1.5
A:Experimental source: strain Bristol N2; clone C42C1
C:Genetics:
A:Gene: CESP:C42C1.5
A:Map position: 1
A:Introns: 1/3; 269/1; 308/3; 343/3
C:Superfamily: mannose-1-phosphate guanylyltransferase
Query Match 56.1%; Score 1045; DB 2; Length 373;
Best Local Similarity 55.3%; Pred. NO. 2.8e-75;
Matches 203; Conservative 64; Mismatches 92; Indels 8; Gaps 3;
QY 1 KRALLVGGGTRLRPLTSLFPKPLVDFAFKPMILHQLKEVGVTEVYVLAIRPEVM 60
Db 9 KRALLVGGGTRLRPLTSLFPKPLVEFAFKPMILHQLKEVGVTEVYVLAIRPEVM 68
QY 61 INFLKDFEDKLGITITCSQETPLTAGPLALARDKADGSGOPFFVLNSDVISEYPPAE 120
Db 69 EQEMTVHADRLGVKLIFSLKEEPLGTAGPLALARKHL-EGDA-PFFVLNSDVICDFFPKQ 126
QY 121 LKFKHCKHGEATIMVTKVDEPSKYGVVYMEATGRVERFVKPTFVGKNKINAGTVLLN 180
Db 127 MYEFHNKHEGTTIATVKVEPSKYGVVVDQDKGIDDFEVEKQYVGVGNKINAGTVIFS 186
QY 181 PSVLDRIELRPTSIEKEVFPQAAADQQLYAMVLPFGFMDVQGPDRDYITGLRLYLDLSIRKK 240
Db 187 SKILDRILPKPTSIEKEIFPEMAFNSLVAFVLPFGFMDVQGPDRDYITGLRLYLDLSIRKK 246
QY 241 SAAKATG-----AHVGNVLVHESAKIGEGCLIGPDVAIGPCVVEDGVRLSRCTVMR 294
Db 247 KSKDLKTSNINHTATIRGNVYDPATYGENCYIGDFYVYVGRVKIEGGVRLHSTILS 306
QY 295 GVRIRKHACISNIGHSTVSGQWARIENMTILGEDVHVDEYVSGGVVLPKHKESSILKP 354
Db 307 DSSIGNYVSGVSGVGRKCHIGSWKRIENICVIGDDVVVKDELILNGASVLPKHSIAVVV 366
QY 355 LRPEIVM 361
Db 367 PSKDIIM 373
RESULT 8
G96778
hypothetical protein F9E10.24 [imported] - *Arabidopsis thaliana*

C:Species: *Arabidopsis thaliana* (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: G96778
R:Rheologis, A.; Ecker, J.R.; Palm, C.J.; Pederspiel, N.A.; Kaul, S.; White, O.; Alon
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.B.; Hughes, R.; Hultzer, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, B.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marzla
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
kei, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Daylis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: G96778
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-415 <STO>
A:Cross-references: GB:AE005173; NID:G6646773; PIDN:AAF21085.1; GSPDB:GN00141
C:Genetics:
A:Gene: F9E10.24
A:Map position: 1
Query Match 29.5%; Score 550; DB 2; Length 415;
Best Local Similarity 31.0%; Pred. NO. 6.1e-36;
Matches 126; Conservative 86; Mismatches 146; Indels 48; Gaps 9;
QY 3 ALIIVGG--FGTRLRPLTSLFPKPLVDFAFKPMILHQLKEV-GVTEVYVLAIRPEV 59
Db 11 AVIMVGQFTGTRFPLSLNPLKPLFFIAQGPVHHFISACKRPNLAQIYLVGFYEER 70
QY 60 MINFLKDFEDKLGITITCSQETPLTAGPLALARDKADGSGOPFFVLNSDVISEYPPAE 119
Db 71 FALVVSATSNELKVPVRYLRDKPHGSAGGLYHFRNLIMEDSPSHILLNCCSPFLP 130
QY 120 ELIKFKHCKHGEATIMVTKV--DEPSKYGVVYMEATGRVERFVKPTFVGKNKINAGTV 177
Db 131 KLEAHRGYGGTITLVIKVSPESASQFELVADPVTNELLHVTEKPTFVSDRINCQV 190
QY 178 LLNPSVL-----DRIELR-----PTS---IEKEVFPQIAADQQLY 209
Db 191 VTFTEIFNAGDVSQTKRDLKRVSSFEALQATPRTIDFVRLDQDQLSPLAGKKRLY 250
QY 210 AMVLPFGFMDVQGPDRDYITGLRLYLDLSIRKKSAKATG-----AHVGNVLVHESAKIG 264
Db 251 TYETMDFEQIKSPGMSLRCSGLYLSQFLTSPOLLASGDGTRSAIVIGDYIHPSAKVH 310
QY 265 EGCILIGPDVAIGPCVVEDGVRLSRCTVMRGVRIKHACISNIGHSTVSGQWARI-- 322
Db 311 PTAKIGPNVSTANARVCGVRLMSCIILDDVEIMENAVVTNAIVGWKSSIGRMSRVOAE 370
QY 323 -----NMVILGEDVHVDEYVSGGVVLPKHKEIKSSILKPEIVM 361
Db 371 GYVNSKLGVTILGDSVAVEDEVVYVTSIVLPNKTILNVSV-QDEIIL 415
RESULT 9
AB2101
mannose-1-phosphate guanylyltransferase [imported] - *Nostoc* sp. (strain PCC 7120)
C:Species: *Nostoc* sp.
A:Note: *Nostoc* sp. Strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120
C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C:Accession: AB2101
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriqu
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing *Cyanobacterium*
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AB2101
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-842 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA074060.1; PID:gl7131453; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: alr2361

Query Match 27.0%; Score 504; DB 2; Length 842;
Best Local Similarity 32.6%; Pred. No. 7.1e-32;
Matches 129; Conservative 83; Mismatches 116; Indels 68; Gaps 16;

QY 1 MKALILVGGFGRTRPLTLSPKPLVDFAFKPMILHQIETALKEVGTETWVLAINRPEVM 60
DB 1 MRVLMAGGSGTRLRPLTCDLPKPWPILNRPPIAHHIINLLKRHHITEVIATLHLPDVL 60
QY 61 INFLKDFEDKLGITITCS-QETPLGTAGPLALARDKLADGSGQPFVLSNDSVISEYPPFA 119
DB 61 RYFQDGDG-FGVQMTVAVEDQPLGTAGCV----KNIAELLDTEFLVISGDSITDFDLG 115
QY 120 ELIKFHKCHGGEATIMTKVDEPSKYGVVVMEEATGRVERFEKRP----KIFVGNKINAGI 176
DB 116 EAIAPFKQKQSKATLILTRVNPPIEFVGVITDEA-GKIKRFLFKPSTSEIF-SDTVNRTG 173
QY 177 YLLNPSVLDRLELPTSTE----KEVFP-QIAADQQLYAMVLPGFWMVDVGQPRDY----I 227
DB 174 YILEPEVL----EYLPSTNTECDFSKDLPLLLAKDEPMPTGYVAEGWCDVGHLDAYREAQY 230
QY 228 TGL-RLYLDGIRKKSAAKATG------AHVGNVNLVHESAKIG------EGCLIGP 271
DB 231 DGLDRKVKLDFARENSPGLWIGQNTYIDPSAHIEAPAVIGNNCRIGARVQIEAGTVIGD 290
QY 272 DVAITGP------GCVVEDGVRLSCTVMRGVRIKKHIA-CISNSIIGHSTVGOWA 319
DB 291 NVITGADANLRKPIVWNGAIIIEEAQLSACVISRGRVDRRAHVLEASVWGSLSVGBEA 350
QY 320 RIENWNTILGEDVHVCDVEYSNGGVVLPHKKEIKSSIL 355
DB 351 QI------SPGVRWPSKKIESGAI 369

RESULT 10

G84459

probable GDP-mannose pyrophosphorylase [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C:Accession: G84459

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon, L.

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: G84459

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-385 <STO>

A:Cross-references: GB:AE002093; NID:g4544432; PIDN:AAD22341.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g04650

A:Map position: 2

Query Match 26.5%; Score 493.5; DB 2; Length 385;
Best Local Similarity 30.2%; Pred. No. 1.7e-31;
Matches 118; Conservative 80; Mismatches 148; Indels 45; Gaps 8;

QY 3 ALILVGG--FGTRLRPLTLSPKPLVDFAFKPMILHQIETALKEV-GVTEVVLAINRPEV 59
DB 8 AVIMVGPTKTRPRPLSFNTPKPLIPLAGPMIHHPIISACKKISLNLAQIIFLIGYEER 67
QY 60 MINFLKDFEDKLGITITCSQTEPLGTAGPLALARDKLADGSGQPFVLSNDSVISEYPPFA 119
DB 68 FALVSSISNELKIPVRYLKDEKPHGSGAGALYFPRDRIMEEKPSHVFLNLCNDCVCSFPLQ 127
QY 120 ELIKFHKCHGGEATIMTKV--DEPSKYGVVVMEEATGRVERFEKRPKIFVGNKINAGIY 177
DB 128 GILDAHRRYGGIGTWLVKVSAAEQSGELIADPDTKELLHYTEKPEFTVSDLLNCGVY 187

QY 178 LLNPSVLDRIE-----LRPTS-----IEKEVFPQIAADQQLYAMVLPGF 216
DB 188 VFTSDIFNAIEEYVSQIRDTSNNYSQATRSVPADFVRLDQDILSLAGKQLYTYENKDF 247
QY 217 WMDVGQPRDYITGLRLYLDSIRKKSAAKATG-----AHVGNVNLVHESAKIGEGCLIG 270
DB 248 WEQIKTPGKSLKCSALYSQFRETSPHILASGDGTNRKPTIIGDVYIHPSVKVLHPTAKIG 307
QY 271 PDVAITPGCVVEDGVRLSCTVMRGVRIKKHACISINSIIGHSTVVGOWARIENWNTILGED 330
DB 308 PNVSISANRVGVGVRLLISCIILDDVEIKENAVVINSIIGKSSIGEAVTVE----- 359
QY 331 VHVCDVSYNGGVVLPHKKEIKSILKPEIVM 361
DB 360 ----DEVAVIGSIVLQNKTLNVSV-QDDIIL 385
RESULT 11
A70363
mannose-1-phosphate guanylyltransferase - Aquifex aeolicus
C:Species: Aquifex aeolicus
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
C:Accession: A70363
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;

V.

Nature 392, 353-358, 1998

A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.

A:Reference number: A70300; MUID:98196666; PMID:9537320

A:Accession: A70363

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-831 <AQF>

A:Cross-references: GB:AE000704; NID:g2983301; PIDN:AAC06893.1; PID:g2983302; GB:AE00

A:Experimental source: strain VF5

C:Genetics:

A:Gene: mpg

Query Match 26.4%; Score 492; DB 2; Length 831;
Best Local Similarity 31.2%; Pred. No. 6.2e-31;
Matches 123; Conservative 86; Mismatches 133; Indels 52; Gaps 14;

QY 1 MKALILVGGFGRTRPLTLSPKPLVDFAFKPMILHQIETALKEVGTETVVLAINRPEVM 60
DB 1 MKGVILAGGFGTRIQIPLTNSIPKPLPVPANRPIMEHVHRLKEAGIEEIVLLYQAEVI 60
QY 61 INFLKDFEDKLGITITCSQTEPLGTAGPLALARDKLADGSGQPFVLSNDSVISEYPPAE 120
DB 61 KNVFKDGS-D-FGVKITVYVQPEADYGTAGAVKQAOYVL----NETFIVSGVITDFNLSE 115
QY 121 LIKFHKCHGGEATIMTKVDEPSKYGVVVMEEATGRVERFEKRP----KIFVGNKINAGIY 177
DB 116 LIAFHKSCKSKFTLALYSVENPLQFGVITNK-EGKVLKELEKPGWGEVF-SDTVNRTG 173

QY 178 LLNPSVLDRI-ELRPTISIEKEVFPQI-AADQQLYAMVLPGFWMVDVGQ------PRDIYT 228
DB 174 VVEPEILNYPEDKPFDMDFAMDLFPKMKSGIDLWALKMRGYWRDIGNIDSYRVDVHKDIFA 233
QY 229 GL------RLYLDGIRKKSAAKATGAHVGNVNLVHESAKIGEG-----CL 268
DB 234 GLVKTRIPGRITTKEARIYVE-----EGTEIPENVSILKGTVIILGNKVKVGESELNKV 288
QY 269 IGPDAVIGPGCVVEDGVRLSCTVMRGVRIKKHACISINSIIGHSTVVGOWARIENWNTILG 328
DB 289 IGNNTVIG------RNKVLFDVSNWNVSIDEESEIRNGVICNDVKIGKRVKAKEGVVIA 342
QY 329 EDVHVCDDEV-YSNGGVVLPHKKEI-KSILKPEIV 360
DB 343 EDCEVEDEVLFKDVVVMPEKVIKESGVVTKNIV 376
RESULT 12
T39403

probable mannose-1-phosphate guanylyl transferase - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 18-Feb-2000
C:Accession: T39403
R:Lyne, M.; Wood, V.; Rajandream, M.A.; Barrell, B.G.; Badcock, K.; Churcher, C.M.
submitted to the EMBL Data Library, April 1998
A:Reference number: 221852
A:Accession: T39403
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-414 <L>N>
A:CROSS-references: EMBL:AL072600; PIDN:CAAL18655.1; GSPDB:GN00067; SPDB:SPBC13G1.02
A:Experimental source: strain 972h; cosmid c13G1
C:Genetics:
A:Gene: SPDB:SPBC13G1.02
A:Map position: 2
A:Introns: 14/2; 27/3
C:Superfamily: mannose-1-phosphate guanylyl transferase

Query Match 26.3%; Score 490.5; DB 2; Length 414;
Best Local Similarity 30.7%; Pred. No. 3.2e-31;
Matches 127; Conservative 71; Mismatches 157; Indels 59; Gaps 10;

QY 3 ALILVGG--FGRRLRLTLSPKPLVDFAFKPMILHQAIEALKEVGVTEVVLAINYRPEVM 60
DB 5 AVILVGGSGRGRFRPLSPKPLVDFAFKPMILHQAIEALKEVGVTEVVLAINYRPEVM 60
QY 61 INFLDKPEKIGI-----TTCSTQETEDICTACPLADKLADGSCGPPFVLNSDVIS 114
DB 63 -SYFKDFINEASHFSPFNRIKYLRYNCLGTGGGLYHFKDQLKHTSNVFMHADVCC 121
QY 115 EYFFAEILRFHCHGGEATIMTKV--DEPSKYGVVVMEEATGRVERFEKPKIFVGNKI 172
DB 122 SPFLOELLNVHKKALVLMATKVKSKEDASNPGCLVEEPTGCVLHVVDKPSLSNII 181
QY 173 NAGIYLLNPSVLDRI-----ELRP-----TSIEKEVFPQIAAD--QOYL 209
DB 182 SCGIYTFDASIFDEIKKAVERRLEEVEKQLRSLDEGMEDVLSLETVDLAPLCSDSKAIY 241
QY 210 AMVLPGFMDVGDOPDIYITGLRLYLD-SIRKSKAAKATCAHVGVNVLVHESAKIGKGL 268
DB 242 AYNTPEFQIKTAGSAVFANSYLQKAYHDGTLKPDTEAEIIPVFIHFNIAIVSKGAK 301
QY 269 IGPDAVAGCCVVEDCVRLSRCTVMRGVRIKHKACISNTHSTHSTVGVQWARIEN----- 323
DB 302 IGPVNSIGARVRIEDGARTNSIIODECEISANAVLHSLSRCKICKWSRVEGCSPTLP 361
QY 324 -----MTLGEDVHVCDEYISNGVGVVPLPHKEIKSSILKPEIYM 361
DB 362 SQHSTTMRNSRVQAITVAGACTIVHDEVRVQNCVLVPHKEIKVGLV-GEIVM 414

RESULT 13
A70978
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: A70978
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: A70978
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-359 <Col>
A:CROSS-references: GB:Z92771; GB:AL123456; NID:G3242259; PIDN:CAB07091.1; PID:gl877319
A:Experimental source: strain H37Rv
C:Genetics:
A:Gene: rmlA2

C:Superfamily: mannose-1-phosphate guanylyl transferase

Query Match 26.2%; Score 488; DB 2; Length 359;
Best Local Similarity 34.6%; Pred. No. 4.2e-31;
Matches 123; Conservative 63; Mismatches 151; Indels 22; Gaps 10;

QY 3 ALILVGGFGRRLRLTLSPKPLVDFAFKPMILHQAIEALKEVGVTEVVLAINYRPEVMIN 62
DB 8 AVILVGGKGRRLRLTLSPKPLVDFAFKPMILHQAIEALKEVGVTEVVLAINYRPEVMIN 64
QY 63 FLKDFED--KLGITITCSQETEPGLTAGPLALARDKLADGGQPFVYVLSNDSVISEYPAE 120
DB 65 FEAEFGGSGALGLGIEVTEBEHPLOTGGGIANVAKLNDTA--MVFNGDVLGADLAQ 121
QY 121 LIKFKHCHGGEATIMTKVDEPSKYGVVVMEEATGRVERFEKPKIFVGNKINAGIYLLN 180
DB 122 LLDHRSNRADYTLQLVRYGDPRAFGCVPTDE-EDRVYVAFLEKTEDPTDQINAGCIYFE 180
QY 181 PSVLDRI-ELRPTSTEKEVFPQIAADQ--QLYAMVLPGFMDVGDOPDRDYITGLRLYLDST 237
DB 181 RNVIDRIPQGRVSVREVERFPDALLADGCKIYGYVDASVYMRDMGTPTDFVRG---SADLV 237
QY 238 RKSKAAKATCAHVGVNVLVHESAKIGKGLIGDPVALGPGCVYVDGVLRSCTVMRGVR 297
DB 238 RGIAPSPALRGHR--GEQLVHDGAASPCALLIGTVVYVGRGAEIAGPGRTRDGAVIDGVR 295
QY 298 IKKHACISNSTIGHSTVGVQWARIENMTI-LGEDVHVCDEVYSNG---GVVLPHKEIKS 352
DB 296 VEAGCVIERSIIGFARIGRPLALIRDGVIGDGADIGARCELLSGARVWPGVFLPDGGIRY 355
QY 353 S 353
DB 356 S 356

RESULT 14
H75070
C:Species: Pyrococcus abyssi
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C:Accession: H75070
R:anonymous, Genoscope
A:Description: Pyrococcus abyssi genome sequence: Insights into archaeal chromosome s
A:Reference number: A75001
A:Accession: H75070
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-413 <KAW>
A:CROSS-references: GB:AJ248286; GB:AL096836; NID:G5458366; PIDN:CAB49069.1; PID:G545
A:Experimental source: Strain Orsay
C:Genetics:
A:Gene: PAB0645
C:Superfamily: glucose-1-phosphate adenyl transferase

Query Match 25.6%; Score 478; DB 2; Length 413;
Best Local Similarity 30.7%; Pred. No. 3.2e-30;
Matches 111; Conservative 87; Mismatches 148; Indels 16; Gaps 6;

QY 1 MKALILVGGFGRRLRLTLSPKPLVDFAFKPMILHQAIEALKEV-GVTEVVLAINYRPEV 59
DB 1 MKAVILAGGFGTRLRPLSPSTRPKPMVPLCKPNQYLQILEALEKYEIDEVILSVYMRGE 60
QY 60 MINFLKDFDKLGIITVCSQETEPGLTAGPLALARDKLADGGQPFVYVLSNDSVISEYPAE 119
DB 61 IREFIQEMRDPKDIRFYNDPMPLETTGGALKNVEEYSD---DFLVIYGDVTFNFDYS 116
QY 120 ELIKFKHCHGGEATIMTKVDEPSKYGVVVMEEATGRVERFEKPKIFVGNKINAGIYLL 179
DB 117 ELIEAHKKNDGLVVAUTKVVDPYDFRFGVITDE-EGKIVEFEKPKPKTKNLVDAGIYMW 175
QY 180 NPSVLDRI-ELRPTSTIEKEVFPQIAADQQLYAMVLPG--FMWDVGPQRDYITGLRLYLD 236

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 27, 2002, 04:07:11 ; Search time 25 Seconds
(without alignments)
598.918 Million cell updates/sec

Title: US-09-374-967-2

Perfect score: 1864

Sequence: 1 MKALILVGGFGRRLRLTSL.....GVVLPKHKSSILKPEIVM 361

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1178	63.2	361	1	MPGL_YEAST
2	339	18.2	355	1	STRD_STRGR
3	287	15.4	257	1	RFBF_SALTY
4	242	13.0	380	1	GLGC_BACSU
5	233.5	12.5	452	1	E2BG_RAT
6	229.5	12.3	452	1	E2BG_HUMAN
7	228	12.2	246	1	SPSI_BACSU
8	227	12.2	456	1	GLMU_ECOLI
9	224.5	12.0	712	1	E2BE_YEAST
10	223.5	12.0	387	1	GLGC_BACST
11	210	11.3	286	1	RFBA_RHISN
12	209	11.2	283	1	YD34_METJA
13	209	11.2	429	1	GLGC_ANASP
14	209	11.2	456	1	GLMU_HAEIN
15	206.5	11.1	459	1	GLMU_BUCAI
16	204	10.9	439	1	GLGC_SYNY3
17	199.5	10.7	516	1	GLGL_MAIZE
18	198	10.6	272	1	YTDA_BACSU
19	197	10.6	431	1	GLGC_SALTY
20	195.5	10.5	293	1	RFPH_ECOLI
21	190	10.2	430	1	GLGC_ECOLI
22	189.5	10.2	297	1	GALE_ECOLI
23	189.5	10.2	297	1	GALE_SALTY
24	189.5	10.2	298	1	GALE_KLEPN
25	184	9.9	289	1	RLMA_STRMU
26	183.5	9.8	250	1	GLGC_BACCL
27	183	9.8	141	1	YSOD_SULAC
28	182.5	9.8	297	1	GALE_SHIFL
29	181.5	9.7	289	1	RFBA_YEREN
30	180.5	9.7	292	1	RFBA_SALAN
31	180.5	9.7	456	1	GCAD_BACSU
32	178	9.5	288	1	RFBA_NEIMB
33	177	9.5	293	1	RBA2_ECOLI

RESULT 1

ID	MPGL_YEAST	STANDARD;	PRT;	361 AA.
AC	P41940;			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	Mannose-1-phosphate guanylttransferase (EC 2.7.7.13) (ATP-mannose-1-phosphate guanylttransferase) (NDP-hexose pyrophosphorylase).			
GN	MPGL OR PSA1 OR YDL055C.			
OS	Saccharomyces cerevisiae (Baker's yeast).			
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;			
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.			
OX	NCBI_TaxID=4932;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=96418871; PubMed=8821656;			
RA	Benton B.K., Plump S.D., Roos J., Lennarz W.J., Cross F.R.;			
RT	"Over-expression of S. cerevisiae G1 cyclins restores the viability of alg1 N-glycosylation mutants.";			
RL	Curr. Genet. 29:106-113(1996).			
[2]				
RP	SEQUENCE FROM N.A.			
RA	Schultz J., Sprague G.F. Jr.;			
RL	Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.			
[3]				
RP	SEQUENCE FROM N.A.			
RA	Bioecker H., Brandt P.;			
RL	Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.			
CC	-!- FUNCTION: ESSENTIAL PROTEIN WHICH MAY BE INVOLVED IN THE REGULATION OF CELL CYCLE PROGRESSION.			
CC	-!- CATALYTIC ACTIVITY: GTP + alpha-D-mannose 1-phosphate -> diphosphate + GDP-mannose.			
CC	-!- SIMILARITY: TO S-TYPHIMURIUM CDP-GLUCOSE PYROPHOSPHORYLASE (RFBF).			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; U19608; AAC49289.1; -			
DR	EMBL; U24437; AAA69677.1; -			
DR	EMBL; Z74103; CAA98617.1; -			
DR	SGD; S0002213; PSA1.			
DR	InterPro; IPR001825; NTP_transferase.			
DR	Pfam; PF00132; hexapep. 4.			
DR	Pfam; PF00483; NTP_transferase; 1.			
KW	Transferase; Kinase; Cell cycle.			
FT	CONFLICT 50 50 V -> A (IN REF. 2).			
SQ	SEQUENCE 361 AA; 39566 MW; DBFIC39BEAE0B776 CRC64;			

Query Match 63.2%; Score 1178; DB 1; Length 361;
Best Local Similarity 61.0%; Pred. No. 1.7e-85;

```
Matches 221; Conservative 56; Mismatches 83; Indels 2; Gaps 2;
QY 1 MKALILVGGFGTRRLRLTLSPFKPLVDFAFKPMILHQIEALKEVGTVVLAIVRPEVM 60
DB 1 MKGLILVGGGTRRLRLTLSPFKPLVDFAFKPMILHQIEALANAGVTDIVLAVNRPV 60
QY 61 INFLKDFEDKLGITTCQETPLGTAGPLALARDKLADGSGOPFFVLSNDVISEYPPAE 120
DB 61 VETLKKYKEYGVNITFVETEPPLGAPGLKLAEDVLKK-DNSPPFVLSNDVICEYDFKE 119
QY 121 LTKPHKCHGGEATIMVTKVDEPSKYGVVWMEAT-GRVERFVEKPKIFVGNKNAGIYLL 179
DB 120 LADFHKHGGKGTIVATKYVDEPSKYGVVHDATPNLIDRFVEKPKFVGNRNAGIYLL 179
QY 180 NPSVLDRIELRSTSEKEVFPQTAADQQLYAMVLPQFMMDVQDRDVIYGLRLYLSIRK 239
DB 180 NPEVDILIMKDTSEKTEFPILVEKQLYSFDLEGFMDVQDPKDFLSGTVLYLSLAK 239
QY 240 KSAALATGAHVGVNVLHESAKIGBGLIGPDVAIGPGCVYEDGGVRLSCTVHGVRIK 299
DB 240 RQPKKLTGAANTVGNALIDPTAKISSTAKIGPDVVTGPNVTIGDGVRIIRSVVLCNSTIK 299
QY 300 KHCISNSTIGWHSVGVQWARTFNMVTLIGEDVHVCDEVTSNGGVLPHPKIKSSILKPEI 359
DB 300 NISLVNSTIYGVNSTVYQWCRLEGVTVLGGDVEVKDEIVKINGKYLPHKISISDNVPKEAT 359
QY 360 VM 361
DB 360 IM 361
RESULT 2
STRD_STRGR STANDARD; PRT; 355 AA.
AC P08075;
DT 01-AUG-1998 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DE Glucose-1-phosphate thymidyltransferase (EC 2.7.7.24) (GTPD-glucose
DE synthase) (GTPD-glucose pyrophosphorylase) (Sugar-nucleotidylolation
DE enzyme).
GN STRD.
OS Streptomyces griseus.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
OX NCBI_taxid=1911;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=N2-3-11;
RX MEDLINE=88040426; PubMed=3118332;
RA Distler J., Ebert A., Mansouri K., Pissowatzki K., Stockmann M.,
RA Pipersberg W.
RT "Gene cluster for streptomycin biosynthesis in Streptomyces griseus:
RT nucleotide sequence of three genes and analysis of transcriptional
RT activity."
RL Nucleic Acids Res. 15:8041-8056(1987).
CC -1- FUNCTION: INVOLVED IN THE BIOSYNTHESIS OF THE STREPTOSE MOIETY OF
CC STREPTOMYCIN.
CC -1- CATALYTIC ACTIVITY: dTTP + alpha-D-glucose 1-phosphate -
CC diphosphate + dTDP-glucose.
CC -1- PATHWAY: Streptomycin biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE GLUCOSE-1-PHOSPHATE
CC THYMIDYLTRANSFERASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL; Y00459; CAA68514.1; --
```

```
DR InterPro: IPR001825; NTP transferase.
DR Pfam: PF00483; NTP transferase; 1.
DR TIGRfams: TIGR01208; tmiA_1ong; 1.
KW Streptomycin biosynthesis; Transferase; Kinase;
KW Nucleotidyltransferase.
SQ SEQUENCE 355 AA; 38098 MW; ED136AEA854EB7DA CRC64;
Query Match 18.2%; Score 339; DB 1; Length 355;
Best Local Similarity 27.9%; Pred. No. 1.7e-19;
Matches 100; Conservative 73; Mismatches 144; Indels 42; Gaps 11;
QY 1 MKALILVGGFGTRRLRLTLSPFKPLVDFAFKPMILHQIEALKEVGTVVLAIVRPEVM 60
DB 1 MKALILVGGGTRRLRLTLSPFKPLVDFAFKPMILHQIEALANAGVTDIVLAVNRPV 60
QY 61 INFLKDFEDKLGITTCQETPLGTAGPLALARDKLADGSGSGOPFFVLSNDVISEYPPAE 114
DB 61 VAAVGD-GSRFGKLVSYIPQSKPLGLAHLYLSRDFLGEDDFIMVLCDFVGVVDSVR 119
QY 115 EYFPAELIKPHKCHGGEATIMVTKVDEPSKYGVVWMEATGRVERFVEKPKIFVGNKN 174
DB 120 EFRAAR-----PDAHLMLTRVPFRSGVYALSD-SQVYLGLEEKPAHPSDLALV 169
QY 175 GIYLLNFSYLDRI-ELRPT-----SIEKEVFPQTAADQQLYAMVLPQFMMDVQDRDVI 228
DB 170 GVYLFSDAIHEAVAAITPSWRGELEITDAVQWLIDAGRDVRSIVISGIMKDTIGNYDML 229
QY 229 GLRLYLSIRKSAKALATGAHVGVNVLHESAKIGBGLIGPDVAIGPGCVYEDGGVRLS 277
DB 230 VNRVLETTBPCDGLYDERSDLIGRVLVECAEVRNRSVMCPVIGAGTRVTSYVGP 289
QY 278 -----GVVEDGVRLSCTVHGVRIKHKACISNIIGWHSVGVQWARTFNMVTLIG 329
DB 290 TSAEDCVWEDS-EVEFSIVLRGASISGVYRRIEASLIGRHVQVTSAPVPHANRLVLAD 347
RESULT 3
RFBF_SALTY STANDARD; PRT; 257 AA.
AC P26336;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DE Glucose-1-phosphate cytidyltransferase (EC 2.7.7.33) (CDP-glucose
DE pyrophosphorylase).
GN RFBF OR STM2092.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2;
RX MEDLINE=91260454; PubMed=1710759;
RA Jiang X.-M., Neal B., Santiago F., Lee S.J., Romana L.K., Reeves P.R.;
RA "Structure and sequence of the rfb (O antigen) gene cluster of
RA Salmonella serovar typhimurium (strain LT2).";
RA Mol. Microbiol. 5:695-713(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dancie M., Du R., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulyanay E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
CC -1- CATALYTIC ACTIVITY: CTP + D-glucose 1-phosphate - diphosphate +
CC CDP-glucose.
CC -1- PATHWAY: CDP-ADENOSE BIOSYNTHESIS WITHIN THE O ANTIGEN
```

"sequencing and functional annotation of the *Bacillus subtilis* genes
 in the 200 kb *rnb-dnaB* region.",
 Microbiology 143:3431-3441(1997).
 [3]
 RT SEQUENCE FROM N.A.
 RN STRAIN=168;
 RC MEDLINE=98044033; PubMed=9384377;
 RX Kunst F., Ogawara N., Mosser I., Albertini A.M., Alloni G.,
 RA Azevedo R., Betero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Borriass R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Bruchi C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.D., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
 RA Guiseppe G., Guy B.J.J., Haga K., Halech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Mostl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Presecan E., Pujic C., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*
 RT *subtilis*,"
 RT Nature 390:249-256(1997).
 CC -I- CATALYTIC ACTIVITY: ATP + alpha-D-glucose 1-phosphate =
 CC diphosphate + ADP-glucose.
 CC -I- PATHWAY: Glycogen biosynthesis; first step.
 CC -I- INDUCTION: EXPRESSED EXCLUSIVELY ON MEDIA CONTAINING CARBON
 CC SOURCES THAT ALLOW EFFICIENT SPOULATION.
 CC -I- SIMILARITY: BELONGS TO THE BACTERIAL AND PLANTS GLUCOSE-1-
 CC PHOSPHATE ADENYLYLTRANSFERASE FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 CC EMBL: Z25795; CAA81041.1; -;
 DR EMBL: AF008220; AAC00215.1; -;
 DR EMBL: Z91119; CAB15075.1; -;
 DR EMBL: S36625; S36625.
 DR PIR: S36625; BG10908; glgC.
 DR Subtilist; BG10908; glgC.
 DR InterPro: IPR001825; NTP_transferase.
 DR Pfam: PF00483; NTP_transferase; 1.
 DR PROSITE: PS00808; ADP_GLC_PYROPHOSPH_1; 1.
 DR PROSITE: PS00809; ADP_GLC_PYROPHOSPH_2; 1.
 DR PROSITE: PS00810; ADP_GLC_PYROPHOSPH_3; 1.
 KW Glycogen biosynthesis; transferase; Nucleotidyltransferase;
 KW Complete proteome.
 KW
 SQ SEQUENCE 380 AA; 42602 MW; 8383169C727A4DB5 CRC64;


```

Db 7 AMIAGGKSGRUSGUTKWKANPAVFGGKYRIIDFTLSNCSNGIDTVGILLTQIQPLELN 66
QY 62 NPLK-----DFED-KLGTIT-----TCSQETBPLATGALALARDKADGSGOPFFVLNS 110
Db 67 SVIGIGSAWLDYNGCVTVLPYAESSEVKWKGTASSIYENLNLNOYDPEYVLISLG 126
QY 111 DYSISYFAELIKFKHKGGEATIMYTKV--DEPSKYGVVYWEATGRVERFEKPKFV 168
Db 127 DHIVRMDYCKMLDYHIERKADVTISVIEGWEEASRFQ-IMKANDPDGTTTHFDKPKPK 185
QY 169 GNKNAGIYLLN-PSVLDRIEL---RPTS---IEKEVFQIARQ-QOLYAMVLPGEFMDV 220
Db 186 SNLASHGRIIFNWPDLKQVLEDDQNPYSYSHDFGKDIIFLLLEKRLKSLAYEPFGYWKDV 245
QY 221 GOPRDYITGLRLYLSIRKSAAKATGCAHVVGCVNLVSHESAKIGECGLICPDVAIGPGCV 280
Db 246 GTVQSL---WEANMDLLKEDSELKUFERKWKIYSVNPQPQFISSDAQVDSLVNEGCV 302
QY 281 VEDGVRLSCTVMRGVRIKKHACISNITGHWSTVGQWARIENMTILGEBDVHVCDEVYSN 340
Db 303 VYGNV--SHSVLFQGVTVGKHTVTSSVIMPDPVTIGEHVVVIAIYPN-----GMVLDP 354
QY 341 GGVLPKHEIKSILKPE 358
Db 355 GAVIRSEKDIIEVLLVSE 372

```

RESULT 5

```

E2BG_RAT STANDARD: PRT: 452 AA.
ID E2BG_RAT
AC F70541;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Translation initiation factor eIF-2B gamma subunit (eIF-2B GDP-GTP
DE Exchange factor).
DE EIF2B3 OR EIF2BG.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Skeletal muscle;
RX MEDLINE=96404916; PubMed=8809057;
RA Price N.T., Kimball S.R., Jefferson L.S., Proud C.G.;
RT Cloning of cDNA for the gamma-subunit of mammalian translation
RT Initiation factor 2B, the guanine nucleotide-exchange factor for
RT eukaryotic initiation factor 2.;
RL Biochem. J. 318:631-636(1996).
CC -!- FUNCTION: CATALYZES THE EXCHANGE OF EUKARYOTIC INITIATION FACTOR
CC -!- 2-BOUND GDP FOR GTP.
CC -!- SUBUNIT: COMPLEX OF FIVE DIFFERENT SUBUNITS; ALPHA, BETA, GAMMA,
CC DELTA AND EPSILON.
CC -!- TISSUE SPECIFICITY: BRAIN, HEART, LIVER, MUSCLE AND TESTES.
CC -!- SIMILARITY: BELONGS TO THE EIF-2B GAMMA/EPSILON SUBUNITS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U38253; AAC52788.1;
DR InterPro: IPR001825; NTP transferase.
DR pfam: PF00483; NTP_transferase; 1
KW Amino-acid biosynthesis; Translation regulation.
SQ SEQUENCE 452 AA; 50407 MW; 18314D39464233F9 CRC64;

```

Query match 12.5%; Score 233.5; DB 1; Length 452;

```

Best Local Similarity 21.5%; Prod No. 4.6e-11;
Matches 96; Conservative 76; Mismatches 144; Indels 137; Gaps 17;

QY 2 KALIVGCGFCFTRLRPLFLSPKPLVDLPANKPMILHQLHQLAEKVGTVVLAINTKPEVMI 61
Db 4 QAVYMAVGGSGRWTDLTSSIFKPLLPVGNKPLIWPENLLERYGFEVIVYT--KDVQK 61
QY 62 NFKLDFEDKLGITTCQETBPLATGALALARDKADGSGOPFFVLNSDVISVPAEL 121
Db 62 ALCAEFKMKLKPDIIVCIPEADMGTSLSRHYPKLA---TDVLVLGCDLITDVALHEV 117
QY 122 IKFHKHGGGEATIMYTKVDE-----PSKYGV-----VMEBATGRVERFV----- 161
Db 118 VDLPRVADSLAMLRKQCGSTEDVPQCKKKKTVFORDFIGVDSTGKRLFLFMANEADLD 177
QY 162 -----EKPKI-FVGNKINAGIYLLNPSVLDRI--ELRPTSIEKEVFPQI----- 202
Db 178 EELVIKGSIIQKPHRIHFQGLVDAHLVCLKKYVVDFLMEKNSITSIRSELPILVLRKQF 237
QY 203 --AADQ-----LYAMV-----LPGF-----WMDVGOPR 224
Db 238 SSASSOQROEDKEDLKKKPKSLDIYSFKKNDTLTLPDYACWNAFRDRKWDLSRSQ 297
QY 225 DYITGLRLYLSIRKSAKAT--GAHVGN-----VLVHESAKIGECGLI 269
Db 298 -----VRCYVHIIMKEGLCSRVSTGLGYMEANRQVPKLLSVLCPEESMIHPSAQIANKHLI 352
QY 270 GPD-----VAIGPGCVVEDGVRLSCTVMRGVRIKKHACISNITGHW 312
Db 353 GADSLGSDQVGEKSSIKRSVIGSSCVIDRVTNVCNLLMNSVTVGEGSSIHGVSIFNN 412
QY 313 STV-----GOWARIE-----NMTILGED 330
Db 413 AVVEAGAEIRDCGLIGSGQRIEAKAKRMNEIVGND 447

RESULT 6
E2BG_HUMAN STANDARD: PRT: 452 AA.
ID E2BG_HUMAN
AC Q9NR50; Q9H850;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Translation initiation factor eIF-2B gamma subunit (eIF-2B GDP-GTP
DE Exchange factor).
DE EIF2B3.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20359747; PubMed=10900014;
RA Krueger M., Beger C., Li Q.-X., Welch P.J., Tritz R., Leavitt M.,
RA Barber J.R., Wong-Staal F.;
RT Identification of eIF2B gamma and eIF2 gamma as cofactors of
RT hepatitis C virus internal ribosome entry site-mediated translation
RT using a functional genomics approach.;
RL PROC. Natl. Acad. Sci. U.S.A. 97:8566-8571(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Ishigaki T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa M., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Wagatsuna M., Hosoliri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA Ninomiya K., Iwayanagi T.;
RP NEDO human cDNA sequencing project.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CATALYZES THE EXCHANGE OF EUKARYOTIC INITIATION FACTOR
CC -!- 2-BOUND GDP FOR GTP.
CC -!- SUBUNIT: COMPLEX OF FIVE DIFFERENT SUBUNITS; ALPHA, BETA, GAMMA,

```



```

Db 59 POFYKLLNGEELGVITYOVQPAASISDGLSYAKRFTKES---FILLGLGNIFEDSL 115
QY 116 YPAEILKHKHCHGGRATIMYTKVDESKYGVVVMEDATGRVERFVKPIFVGNKINAG 175
Db 116 KPYTE--RFEQ-QCKGAKVLLKVEDDEREGIAIDEKKNRIRSIIEKPEOPTNLCVTG 172
QY 176 IYLLNSVLDRIELRPTISIEKEVPQTAADQOLYAM-----VLPGFWMVGGPRDYIT 228
Db 173 IYNDAEVFSYIEQISPSKRGEL--EITDNNLYIENSQLTVDYLSGWWVDAG-----T 224
QY 229 GLRLVLDSTIRKSAAKLATCAHVGVNVLVHESAKIGE 265
Db 225 HESLYLAS-----QLVHQALNKGQ 243

RESULT 8
GLMU_ECOLI STANDARD; PRT; 456 AA.
AC P17114; P76746;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Bifunctional glmu protein [includes: UDP-N-acetylglucosamine
DE pyrophosphorylase (EC 2.7.7.23) (N-acetylglucosamine-1-phosphate
DE uridylyltransferase); Glucosamine-1-phosphate N-acetyltransferase
DE (EC 2.3.1.57)].
GN GLMU OR B3730 OR 25228 OR EGS4672.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TAXID=562, 83334;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85121806; PubMed=6395859;
RA Walker J.E., Gay N.J., Saraste M., Eberle A.N.;
RT "DNA sequence around the Escherichia coli unc operon. Completion of
RT the sequence of a 17 kilobase segment containing asuA, oric, unc,
RT glms and phos.";
RL Biochem. J. 224:799-815(1984).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=93315143; PubMed=7686082;
RA Burland V.D., Plunkett G. III, Daniels D.L., Blattner F.R.;
RT "DNA sequence and analysis of 136 kilobases of the Escherichia coli
RT genome: organizational symmetry around the origin of replication.";
RL Genomics 16:551-561(1993).
RN [3]
RP REVISIONS.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans D.S., Gregor J., Kirkpatrick H.A.,
RA Postai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoudis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohemorrhagic Escherichia coli O157:H7.";
RL Nature 409:528-533(2001).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;

```

```

RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsuda E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kihara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
RN [6]
RP IDENTIFICATION.
RX MEDLINE=94012475; PubMed=9407787;
RA Mengin-Lecreulx D., van Heijenoort J.;
RT "Identification of the glmu gene encoding N-acetylglucosamine-1-
RT phosphate uridylyltransferase in Escherichia coli.";
RL J. Bacteriol. 175:6150-6157(1993).
RN [7]
RP CHARACTERIZATION.
RX MEDLINE=94364959; PubMed=9083170;
RA Mengin-Lecreulx D., van Heijenoort J.;
RT "Purification of glucosamine-1-phosphate acetyltransferase and
RT N-acetylglucosamine-1-phosphate uridylyltransferase activities of
RT Escherichia coli: Characterization of the glmu gene product as a
RT bifunctional enzyme catalyzing two subsequent steps in the pathway
RT for UDP-N-acetylglucosamine synthesis.";
RL J. Bacteriol. 176:5788-5795(1994).
RN [8]
RP CHARACTERIZATION.
RX MEDLINE=96140233; PubMed=8555230;
RA Gehring A.M., Lees W.J., Mindiola D.J., Walsh C.T., Brown E.D.;
RT "Acetyltransfer precedes uridylyltransfer in the formation of
RT UDP-N-acetylglucosamine in separable active sites of the bifunctional
RT Glmu protein of Escherichia coli.";
RL Biochemistry 35:579-585(1996).
CC -1- FUNCTION: BIFUNCTIONAL ENZYME RESPONSIBLE FOR THE ACETYLATION OF
CC GLC-N-1-P TO GIVE GLCNAC-1-P AND THE SYNTHESIS OF UDP-GLCNAC.
CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + D-glucosamine 1-phosphate = CoA +
CC N-acetyl-D-glucosamine 1-phosphate.
CC -1- CATALYTIC ACTIVITY: UTP + N-acetyl-alpha-D-glucosamine
CC -1- phosphate + UDP-N-acetyl-D-glucosamine.
CC -1- PATHWAY: PEPTIDOGLYCAN BIOSYNTHESIS; LIPOPOLYSACCHARIDE
CC BIOSYNTHESIS (LIPID A BIOSYNTHESIS).
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE
CC CYSE/LACA/LPXA/NOFL FAMILY OF ACETYLTRANSFERASES. COMPOSED OF
CC MULTIPLE REPEATS OF [LIV]-G-X(4).
CC -1- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC FRAMESHIFT THAT CREATES TWO ORFS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X01631; CAA25784.1; -
DR EMBL; L10328; AAA62082.1; ALT_FRAME.
DR EMBL; L10328; AAA62081.1; ALT_FRAME.
DR EMBL; AE000450; AAC76753.1; -
DR EMBL; AE005605; AAC58933.1; -
DR EMBL; AP002566; BAB38095.1; -
DR EMBL; EG11198; glmu.
DR EMBL; IP001451; Hexapep.transf.
DR InterPro; IPR001825; NTP_transferase.
DR Pfam; PF00132; hexapep; 7.
DR Pfam; PF00483; NTP_transferase; 1.
DR TIGRfam; TIGR01173; glmu; 1.
DR PROSITE; PS00101; HEXAPEP_TRANSFERASES; 1.
KW Transferase; Acyltransferase; Nucleotidyltransferase;
KW Multifunctional enzyme; Peptidoglycan synthesis; Cell wall; Repeat;
KW Complete proteome.
FT CONFLICT 186 187
SQ SEQUENCE 456 AA; 40190 MW; 2P3C5C84P673C8A3 CRC64;

```

```
Query Match 12.2%; Score 227; DB 1; Length 456;
Best Local Similarity 24.4%; Pred. No. 1.5e-10;
Matches 100; Conservative 55; Mismatches 158; Indels 96; Gaps 16;

QY 1 MKALLIVGGTGRRLPLTLSPFPKPLVDFAKPMILHQLAEALKEVGVTEVVLAINYRPEVM 60
DB 6 MSVILAAGKGRMYS---DLPKVLHTLAGKAMVQHVDAANELGAHVHLVYGHGGDLL 62
QY 61 INFLKDFEDKIGITTCSETEPLGTAGPLALARDKLADSGSGPFLVNSDVISEYPPFAE 120
DB 63 KOALAD--DNLNWL-----QAEQLGTGHAMQQA-----APFFADDEDILMLYGDVP 107
QY 121 LIKFH-----KCHGEATIMVTKVDEPSKYGVVME--EATGRVER---FVEKPKIF 167
DB 108 LISVETLQRLRDAPKQGG-IGLLTVKLDPTGCGRTRENGKVTGIVEHKDATDEORQI- 165
QY 168 VGNKINAGIYLLN-----PSVLDRI 187
DB 166 --QEINTGILANGAMKRWLAKLTNNNAQGEYITDIILALAYQEGREIVAVHPQRLSEV 223
QY 188 ELRPSIEKEVPPQTAADQOLYAMVLPFGFWMVDVGQPRDYIT-GLRLYLDLSIRKKSAAKLA 246
DB 224 EGVNRLQLSLERYVQSEAEKLLAGVYL-----RDPAREDLRLGTLTHGRD---VEID 275
QY 247 TGAHVGVNVLVHESAKIGECG-----LIGPDVAIGPGCVVDEGVRLSRCTV-----MRG 295
DB 276 TNVIEGNTVTLGHRVKGITGCVKNSVIGDDCEISFYTVVEDANAACTIGPFARLRPG 335
QY 296 VRIKHACISNIIIGWSTVGQWARIENMTILGEDVHVGVNSGVV 344
DB 336 AELLEGAVHGVNFEMKARLGRKSGAGHLYLG-DAEIGDNVNIAGTI 383

RESULT 9
ID EB2E_YEAST STANDARD; PRT; 712 AA.
AC P32501;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Translation initiation factor eIF-2B epsilon subunit (eIF-2B GDP-GTP
DE exchange factor) (Guanine nucleotide exchange factor subunit GCD6)
DE (GCD complex subunit GCD6).
GN GCD6 OR YIF225 OR YDR211W OR YD8142.12 OR YD8142B.03.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93180841; PubMed=8441423;
RA Bushman J.L., Asuru A.I., Matts R.L., Hinnebusch A.G.;
RT "Evidence that GCD6 and GCD7, translational regulators of GCN4, are
RT subunits of the guanine nucleotide exchange factor for eIF-2 in
RT Saccharomyces cerevisiae."
RL Mol. Cell. Biol. 13:1920-1932(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RA Oliver K., Shore L., Harris D., Barrell B.G., Rajandream M.A.,
RA Walsh S.V.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: SUBUNIT OF THE GUANINE NUCLEOTIDE EXCHANGE FACTOR FOR
CC EIF-2. REQUIRED TO REPRESS GCN4 TRANSLATION UNDER NONSTARVATION
CC CONDITIONS. GCD6 AND GCD7 REPRESS GCN4 EXPRESSION AT THE
CC TRANSLATIONAL LEVEL BY ENSURING THAT RIBOSOMES WHICH HAVE
CC TRANSLATED UORF1 WILL REINITIATE AT UORF2, -3, OR -4 AND THUS FAIL
CC TO REACH THE GCN4 START SITE.
CC -1- SUBUNIT: COMPLEX OF FIVE DIFFERENT SUBUNITS, ALPHA (GCN3), BETA
CC (GCD7), GAMMA (GCD1), DELTA (GCD2) AND EPSILON (GCD6).
CC -1- SIMILARITY: BELONGS TO THE EIF-2B GAMMA/EPSILON SUBUNITS FAMILY.
-----
```

```
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; L07115; AAA65498.1; -
DR EMBL; Z68194; CAA92354.1; -
DR EMBL; Z68195; CAA92362.1; -
DR PIR; S30776; S30776.
DR PIR; A48156; A48156.
DR SGD; S0002619; GCD6.
DR InterPro; IPR001451; Hexapep.transf.
DR InterPro; IPR003307; eIF5C.
DR Pfam; PF00132; hexapep; 4.
DR Pfam; PF02020; W2; 1.
DR SMART; SM00515; eIF5C; 1.
KW Amino-acid biosynthesis; Translation regulation.
SQ SEQUENCE 712 AA; 81161 MW; EFE87F6AE2941619 CRC64;

Query Match 12.0%; Score 224.5; DB 1; Length 712;
Best Local Similarity 22.3%; Pred. No. 4.1e-10;
Matches 92; Conservative 76; Mismatches 155; Indels 89; Gaps 12;

QY 1 MKALLIVGGTGRRLPLTLSPFPKPLVDFAKPMILHQLAEALKEVGVTEVVLAINYRPEVM 60
DB 27 LQAVLTDSYETRFMPLTAVKPRCLLPANVPLIETLEFLAKAGVHEVFLICSHANQI 86
QY 61 INFLKDFEDKIL-----GITTCSETEPLGTAGPLALARDKLADSG---QPFVFLNSDV 112
DB 87 NDYIENSKWNLWSPFPKITTMSPEARCTGD-----VARD--LDNRGIITGDFILVSGDV 139
QY 113 ISEYPAELIKFKH-----CHGEATIMVTKVDEPSKYGVVMEETATGRVER 159
DB 140 LTNIDFSKMLEFHKHMLQDKDHTWCLSKASTYPTKRTIEPAF---VLDKSTSRCY 196
QY 160 FVEKPKIFVGNKINAGIYLLNPSVLDRIE----- 188
DB 197 YQDLPLPSSREKTSIQI---DPELLDNVDFVIRNDLIDRICTSHVPLIFQENFDYQ 253
QY 189 -LRPTSTEKEVFPQIAADQOLYAMVLPFGFWMVDVGQPRDYITGLR-----LYLDSIR 238
DB 254 SLR-TDFVKGVISSDILGKHIIAYLTDEYAVRVESWQTYDTISQDFLGRWCYPLVLDNSI 312
QY 239 KSAAKLATGAHV--GNVLVHESAKIGEGCLIGPDVAIGPGCVVDEGVRLSRCTVMRGV 296
DB 313 QDDQTVSYESRHIYKEDVVLQAQCKIGKCTAIGSGTKIGESTKIENSVIGRNCQIGENI 372
QY 297 RIK-----KHACISNIIIGWSTVGQWARIENMTILGEDVHVGVNSGVV 337
DB 373 RIKNSFIWDDCIIGNSIIDHSLIASNATLGSNVRLNDGCIIGPNVKIDNNM 424

RESULT 10
GLGC_BACST STANDARD; PRT; 387 AA.
ID GLGC_BACST
AC O08326;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Glucose-1-phosphate adenylyltransferase (EC 2.7.7.27) (ADP-glucose
DE synthase) (ADP-glucose pyrophosphorylase).
GN GLGC.
OS Bacillus stearothermophilus.
OC Bacteria; Firmicutes; Bacilliales; Geobacillus.
OX NCBI_TaxID=1422;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=TRBE14;
RX MEDLINE=97386405; PubMed=9244254;
RA Takata H., Takana T., Okada S., Takagi M., Imanaka T.;
```

"Characterization of a gene cluster for glycogen biosynthesis and a heterotetrameric ADP-glucose pyrophosphorylase from *Bacillus stearothermophilus*."
 J. Bacteriol. 179:4689-4698(1997).
 -1- CATALYTIC ACTIVITY: ATP + alpha-D-glucose 1-phosphate -> diphosphate + ADP-glucose.
 -1- PATHWAY: Glycogen biosynthesis; first step.
 -1- SIMILARITY: BELONGS TO THE BACTERIAL AND PLANTS GLUCOSE-1-PHOSPHATE ADENYLYLTRANSFERASE FAMILY.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 EMBL: D87026; BAA19589.1; -
 InterPro: IPR001825; NTP_transferase.
 Pfam: PF00483; NTP_transferase; 1.
 PROSITE: PS00808; ADP_GLC_PYROPHOSPH_1; 1.
 PROSITE: PS00809; ADP_GLC_PYROPHOSPH_2; 1.
 PROSITE: PS00810; ADP_GLC_PYROPHOSPH_3; 1.
 Glycogen biosynthesis; Transferase; Nucleotidyltransferase.
 Glycogen biosynthesis; Transferase; 3660DB962442D97B CRC64;
 SEQUENCE 387 AA; 43268 MW; 3660DB962442D97B CRC64;

 Query Match 12.0%; Score 223.5; DB 1; Length 387;
 Best Local Similarity 23.8%; Pred. No. 2.3e-10;
 Matches 90; Conservative 64; Mismatches 157; Indels 67; Gaps 14;

 QY 3 ALLIVGGFGRRLRLPLTLSPFKPLVDFAKPMIL-HOTEALKEVGVTVEVLAINRPEVMI 61
 DB 7 AMLLAGGQSGKRLSLTNIAPKPAVFGGKYRIIDFTLSNCTNSGIDTVGLVTQIOPLLH 66
 QY 62 NFLK-----DFEDKL-GITITCSQETP-----LGTAGPLALARDKLDGSGGQFFVLNS 110
 DB 67 SYIGISAWDLDRNGGVTVPYSSVSSGVKWEVTANAVQNTNYTEQVNDPVDVLISG 126
 QY 111 DVISEYFPAELIRFKHCHGGEATIMTKV--DEPSKYGVVMEATGRVERFEVKPKIFV 168
 DB 127 DHIYKMDYQHMIDYHTAKQADVTSIVTEWEASRGIMNTNEMEIVE-FAEKPAEPK 185
 QY 169 GNKINAGIYLLNPSVLDRI----ELRPTS---IEKEVFPQTAADQOL-YAMVLPCFQMDV 220
 DB 186 SNLASMGIYFNPLKQYLQIDNANPHSHDFGKVDIPIMLLREKKRPFAYPFEGYKDV 245
 QY 221 GQPRD-----YITGLRLYLDSTRKSAKLAATGCAHVGVNVLVHESAKI 263
 DB 246 GTVKSLWEANMDDLNDENNELDLDRSWRIY--SYNPNOPQIYSPEAEVSDSLVNE---- 299
 QY 264 GEGCLIGPDVAIGFCYVEDSGYRLSCTYMRGVRKIKKACISNLSLIGHSTVGOMARLEN 323
 DB 300 -----GCVVEGTV--ERSVLFQGVGRIGKAVKVESVIMPGAAYSEGAYVER 343
 QY 324 MTILGEDV-----HVCDE 336
 DB 344 AIVTPDSIIIPPHSSYCFE 361

 RESULT 11
 ID RFBA_RHISN STANDARD; PRT; 286 AA.
 AC P55464;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Probable glucose-1-phosphate thymidyltransferase (EC 2.7.7.24)
 DE (dUDP-glucose synthase) (dUDP-glucose pyrophosphorylase).
 GN Y4GH.
 OS Rhizobium sp. (strain NGR234).
 OC Plasmid sym pNGR234a.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

Rhizobiaceae; Rhizobium.
 NCBI_TaxID=394;
 [1]
 SEQUENCE FROM N.A.
 MEDLINE-97303956; PubMed-9163424;
 Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A., Perret X.;
 "Molecular basis of symbiosis between *Rhizobium* and legumes."
 Nature 387:394-401(1997).
 -1- CATALYTIC ACTIVITY: dTTP + alpha-D-glucose 1-phosphate -> diphosphate + dTDP-glucose.
 -1- PATHWAY: DTPP-L-RHAMNOSE BIOSYNTHESIS
 -1- SIMILARITY: BELONGS TO THE GLUCOSE-1-PHOSPHATE THYMIDYLYLTRANSFERASE FAMILY.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 EMBL: AE000074; AAB91682.1; -
 InterPro: IPR001825; NTP_transferase.
 Pfam: PF00483; NTP_transferase; 1.
 TIGRFAMS: TIGR01207; tmlA; 1.
 Hypothetical protein; Plasmid; Transferase; Kinase;
 Nucleotidyltransferase.
 SEQUENCE 286 AA; 31205 MW; 2FD7BFF7575914F2 CRC64;

 Query Match 11.3%; Score 210; DB 1; Length 286;
 Best Local Similarity 25.8%; Pred. No. 1.9e-09;
 Matches 72; Conservative 54; Mismatches 131; Indels 22; Gaps 9;

 QY 1 MKALLVGGFGRRLRLPLTLSPFKPLVDFAKPMILHQIEALKEVGVTVEVLAINRPEVM 60
 DB 1 MKGIIAGSGSTRLPMTLVMSKQILPVYDKPMIPVPLTTLMLAGIRILLIISTPHHML 60
 QY 61 INFLKDFEDKDGITTCSETEPLGTAGPLALARDKLDGSGOP--FFVLNSDVISEYFPA 119
 DB 61 FQALLGDSGQSGISIRYAVQSPNGLAQYVIGADFVA---GSPSCLLDGNDIYFGHGLQ 117
 QY 120 ELIKPHKHGGEATIMTKVDEPSKYGVVMEATGRVERFEVKPKIFVGNKINAGIYLL 179
 DB 118 GILQQAALQOGATIFAYHVDPERYGVVEFDEGNALS-IEEKPAAPKSTWAVTGLIFY 176
 QY 180 NPSVLD-RIELRPTSIEKEVFPQTAADQOLY-----AMVLPGF-HMDVGGQPRDYTG 229
 DB 177 DSEVVDTAANLKPSARCEV---EITDVNRIVLERCKLVAVILGRGVAMLDGTP-DSLLE 232
 QY 230 LRLYLDSTRKSAKLAATGCAHV---VGNVLVHESAKIGE 265
 DB 233 AAEVSTLEKRGKQAFKVCPEEVALANGFISVEEFARIAE 271

 RESULT 12
 ID YD34_METJA STANDARD; PRT; 283 AA.
 AC Q58730;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Putative UTP-glucose-1-phosphate uridylyltransferase (EC 2.7.7.9)
 DE (UDP-glucose pyrophosphorylase) (UDPGP) (Alpha-D-glucosyl-1-phosphate uridylyltransferase) (Uridine diphosphoglucose pyrophosphorylase).
 GN MJ1334
 OS Methanococcus jannaschii.
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
 OC Methanocaldococcaceae; Methanocaldococcus.
 NCBI_TaxID=2190;
 [1]
 SEQUENCE FROM N.A.


```

QY 337 YTSNGGVYLPFKHKKSSILKPEIV 360
| | | | |
DB 408 FVTRSGIVV-----VLKNAVI 423

RESULT 14
GLMU_HAEIN STANDARD; PRT; 456 AA.
AC P43889;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Bifunctional glmu protein [Includes: UDP-N-acetylglucosamine
DE pyrophosphorylase (EC 2.7.7.23) (N-acetylglucosamine-1-phosphate
DE uridylyltransferase); Glucosamine-1-phosphate N-acetyltransferase
DE (EC 2.3.1.57)].
GN GLMU OR H10642.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RD / KW20 / ATCC 51907;
RA FLEISCHMANN R.D., ADAMS M.D., WHITE O., CLAYTON R.A., KIRKNESS E.F.,
RA MCKENLAY A.R., BULT C.J., TOMB J.-F., DOUGHERTY B.A., MERRICK J.M.,
RA MCKENNEY K., SUTTON G., FITZHUGH W., FIELDS C.A., GOCAYNE J.D.,
RA SCOTT J.D., SHIRLEY R., LIU L.-F., GLODEK A., KELLEY J.M.,
RA WEIDMAN J.F., PHILLIPS C.A., SPRIGGS T., HEDBLUM E., COTTON M.D.,
RA UTTERBACK T.R., HANNA M.C., NGUYEN D.T., SAUDEK D.M., BRANDON R.C.,
RA FINE L.D., FRITZCHMAN J.L., FUHRMANN J.L., GEORGHAGEN N.S.M.,
RA GNEHM C.L., McDONALD L.A., SMALL K.V., FRASER C.M., SMITH H.O.,
RA VENTER J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd."
RL Science 269:496-512(1995).
CC -!- FUNCTION: BIFUNCTIONAL ENZYME RESPONSIBLE FOR THE ACETYLYATION OF
CC GLC-N-1-P TO GIVE GLCNAC-1-P AND THE SYNTHESIS OF UDP-GLCNAC
CC (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: Acetyl-CoA + D-glucosamine 1-phosphate = CoA +
CC N-acetyl-D-glucosamine 1-phosphate.
CC -!- CATALYTIC ACTIVITY: UTP + N-acetyl-alpha-D-glucosamine 1-phosphate
CC = diphosphate + UDP-N-acetyl-D-glucosamine.
CC -!- PATHWAY: PEPTIDOGLYCAN AND LIPOPOLYSACCHARIDE BIOSYNTHESIS.
CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE
CC CYSE/LACA/LPXA/NODL FAMILY OF ACETYLTRANSFERASES. COMPOSED OF
CC MULTIPLE REPEATS OF [LIV]-G-X(4).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U32747; AAC22302.1; -
CC TIGR; H10642; -
CC InterPro; IPR001451; Hexapep transf.
CC InterPro; IPR001825; NTP transferase.
CC Pfam; PF00132; hexapep; 6.
CC Pfam; PF00483; NTP transferase; 1.
CC TIGRfams; TIGR01173; glmu; 1.
CC PROSITE; PS00101; HEXAPEP-TRANSFERASES; 1.
KW Transferrase: Acyltransferase; Nucleotidyltransferase;
KW Multifunctional enzyme; Peptidoglycan synthesis; Cell wall; repeat;
KW Complete proteome.
KW SEQUENCE 456 AA; 49287 MW; 93B76532A9F9BD36 CRC64;
Query Match 11.2%; Score 209; DB 1; Length 456;
Best Local Similarity 23.5%; Pred. No. 4e-09;

```

```

Matches 95; Conservative 63; Mismatches 159; Indels 88; Gaps 15;
QY 1 MKALILVCGCTRLRLPLTLSPKPLDPRANKPMILHQTALKEVGYTYVYLAIRPBYM 60
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 6 LSAVILAAGKGRMYS---DLPKVLHTTAGKPMVKHVIDTAHQLGSENIHLIVHGGDLM 62
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 61 INFLKDFEDKLGITTCSTOFTDLCTAGPLALADKLDAGSCOPFFVLNSDVISRYPAE 120
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 63 RTHLANEO-----VNVVLQTEOUGTAHAYOOA-----AFFKKNENIVVLYGAP 107
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 121 LIKEH-----KCHGEATIMYTKVDEPSKYGVVMEETATGRVERVE----- 162
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 108 LITRETFLEKLEAPDENCIA-LLTVNLNDPFCYGRLINE--NGNVVAIVEQKDAEQLN 164
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 163 -----KPKIFYGNKINAGIYLLNF--SYLDRIELRPTSTIE-REVFPQI 202
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 165 IKEVNTGVVYSDGASFKKRLARVGNNAQGEVLTDLTALANQNCQVAVQATDMEVE 224
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 203 AADQOLYAVLPGFMVQCPDPRDVTGLRLYLDSTRKSSAAKLATGAV-----VGN 254
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 225 GANNRLOALALERYFONKQASKLLEGVMIY-DFARFDRLGTLEHGKDYEDVNVYIEGN 283
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 255 VLVHESAKIGESC-----LIGPDVAIGPGCVVDEGVRLSRTV-----MRGVRIKKHAC 303
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 284 VKLGRVKGICGCVLKNVVICNDVEIKDYSVLEDSIVGEKAAIGFFSLRPGAEALAEETH 343
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 304 ISNSIIGHWSTVGOWARLENMTILGEDVHVGDEYISN-----GGVY 344
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 344 VGNFVEIKKSTVGKSKVNHLYTVGD-----SEIGSNICNIGAVI 383
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 15
GLMU_RUCAI STANDARD; PRT; 459 AA.
AC P57139;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Bifunctional glmu protein [Includes: UDP-N-acetylglucosamine
DE pyrophosphorylase (EC 2.7.7.23) (N-acetylglucosamine-1-phosphate
DE uridylyltransferase); Glucosamine-1-phosphate N-acetyltransferase
DE (EC 2.3.1.57)].
GN GLMU OR BU037.
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
OS symbiotic bacterium).
OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OX NCBI_TaxID=118099;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tokyo 1998;
RX MEDLINE=20445173; PubMed=10993077;
RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.,
RT "Genome sequence of the endocellular bacterial symbiont of aphids
RT Buchnera sp. APS."
RL Nature 407:81-86(2000).
CC -!- FUNCTION: BIFUNCTIONAL ENZYME RESPONSIBLE FOR THE ACETYLYATION OF
CC GLC-N-1-P TO GIVE GLCNAC-1-P AND THE SYNTHESIS OF UDP-GLCNAC
CC (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: Acetyl-CoA + D-glucosamine 1-phosphate = CoA +
CC N-acetyl-D-glucosamine 1-phosphate.
CC -!- CATALYTIC ACTIVITY: UTP + N-acetyl-alpha-D-glucosamine 1-phosphate
CC = diphosphate + UDP-N-acetyl-D-glucosamine.
CC -!- PATHWAY: PEPTIDOGLYCAN AND LIPOPOLYSACCHARIDE BIOSYNTHESIS.
CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE
CC CYSE/LACA/LPXA/NODL FAMILY OF ACETYLTRANSFERASES. COMPOSED OF
CC MULTIPLE REPEATS OF [LIV]-G-X(4).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 27, 2002, 05:01:58 ; Search time 82 Seconds
(without alignments)
907.110 Million cell updates/sec

Title: US-09-374-967-2

Perfect score: 1864

Sequence: 1 MKALILVGGFGRRLRLTLTSL.....GVVLPHKEIKSILKPEIVM 361

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*

- 1: sp.archaea.*
- 2: sp.bacteria.*
- 3: sp.fungi.*
- 4: sp.human.*
- 5: sp.invertebrate.*
- 6: sp.mammal.*
- 7: sp.mhc.*
- 8: sp.organelle.*
- 9: sp.phage.*
- 10: sp.plant.*
- 11: sp.rodent.*
- 12: sp.virus.*
- 13: sp.vertibrate.*
- 14: sp.unclassified.*
- 15: sp.virus.*
- 16: sp.bacteriap.*
- 17: sp.archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1766	94.7	361	10 Q941T9	Q941T9 oryza sativ
2	1681	90.2	361	10 Q22287	Q22287 arabidopsis
3	1680	90.1	361	10 Q92TW5	Q92TW5 solanum tub
4	1674	89.8	361	10 Q9C5B8	Q9C5B8 arabidopsis
5	1671	89.6	361	10 Q94IA7	Q94IA7 nicotiana t
6	1587.5	85.2	364	10 Q9M2S0	Q9M2S0 arabidopsis
7	1376	73.8	351	10 Q9M0A3	Q9M0A3 arabidopsis
8	1183.5	63.5	360	4 Q9Y5P6	Q9Y5P6 homo sapien
9	1176.5	63.1	364	3 Q74624	Q74624 trichoderma
10	1175	63.0	363	3 Q74484	Q74484 schizosacch
11	1173	62.9	369	5 Q9VN61	Q9VN61 drosophila
12	1170.5	62.8	364	3 Q9P8N0	Q9P8N0 pichia angu
13	1160	62.2	387	4 Q9H7U3	Q9H7U3 homo sapien
14	1121	60.1	361	3 Q9Y725	Q9Y725 candida gla
15	1099.5	59.0	362	3 Q93827	Q93827 candida alb
16	1045	56.1	373	5 Q44970	Q44970 caenorhabdi

17	967	51.9	379	5 Q9U1C2	Q9U1C2 leishmania
18	953	51.1	379	5 Q9BLW4	Q9BLW4 leishmania
19	789.5	42.4	345	5 Q8SQX7	Q8SQX7 encephalito
20	594.5	31.9	420	11 Q922H4	Q922H4 mus musculu
21	592.5	31.8	420	4 Q9NWC3	Q9NWC3 homo sapien
22	592.5	31.8	420	4 Q96IJ6	Q96IJ6 homo sapien
23	551.5	29.6	399	4 Q9Y5P5	Q9Y5P5 homo sapien
24	550	29.5	411	10 Q9SSG7	Q9SSG7 arabidopsis
25	550	29.5	415	10 Q9C9P3	Q9C9P3 arabidopsis
26	539	28.9	401	5 Q9N4V3	Q9N4V3 caenorhabdi
27	536	28.8	406	10 Q8W4J5	Q8W4J5 arabidopsis
28	523	28.1	438	5 Q8SXU3	Q8SXU3 drosophila
29	521	28.0	448	5 Q9V7E1	Q9V7E1 drosophila
30	514	27.6	394	5 Q9N4V2	Q9N4V2 caenorhabdi
31	509	27.3	831	16 Q9K2P4	Q9K2P4 streptomyce
32	504	27.0	842	16 Q8YU15	Q8YU15 anabaena sp
33	493.5	26.5	385	10 Q9S3J4	Q9S3J4 arabidopsis
34	492	26.4	831	16 Q66933	Q66933 aquifex aeo
35	490.5	26.3	414	3 Q60064	Q60064 schizosacch
36	489	26.2	392	17 Q8TL99	Q8TL99 methanosarc
37	488	26.2	359	16 Q96869	Q96869 mycobacteri
38	483	25.9	413	17 Q8U2G7	Q8U2G7 pyrococcus
39	482	25.9	360	16 Q9KZK6	Q9KZK6 streptomyce
40	480	25.8	778	16 Q9R8I4	Q9R8I4 thermoanaer
41	478	25.6	413	17 Q9V037	Q9V037 pyrococcus
42	475	25.5	416	17 Q58775	Q58775 pyrococcus
43	473.5	25.4	361	17 Q980H9	Q980H9 sulfolobus
44	467.5	25.1	357	17 Q96Y83	Q96Y83 sulfolobus
45	467	25.1	349	16 Q8RDG7	Q8RDG7 thermoanaer

ALIGNMENTS

RESULT 1

Q941T9 ID Q941T9 PRELIMINARY; PRT; 361 AA.
AC Q941T9;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Putative GDP-mannose pyrophosphorylase.
GN P0005H10.9 OR P0446B05.26.
OS Oryza sativa (Rice), and
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530, 39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
clone:P0005H10."
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
clone:P0446B05."
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP004127; BAB64272.1; -.
DR EMBL; AP003251; BAB89577.1; -.
DR InterPro; IPR001451; Hexapep-transf.
DR InterPro; IPR001825; NTP-transferase.
DR Pfam; PF00132; hexapep; 4.
DR Pfam; PF00483; NTP-transferase; 1.
DR PROSITE; PS00101; HEXAPEP_TRANSFERASES; UNKNOWN_1.
SQ SEQUENCE 361 AA; 39606 MW; 9E0B29A8107075D2 CRC64;

Query Match 94.7%; Score 1766; DB 10; Length 361;

```

Best Local Similarity 93.6%; Pred. No. 1.4e-136;
Matches 338; Conservative 14; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKALLVGGFGTGLRPLTSLSPKPLVDFAFKPMILHQIEALKEVGTVTVLAINRPEVM 60
Db 1 MKALLVGGFGTGLRPLTSLSPKPLVDFAFKPMILHQIEALKEVGTVTVLAINRPEVM 60

QY 61 INFLKDFEDKLGITTCSETPLGTAGPLALARDKLDGSGQFFVLNSDVISEYFAE 120
Db 61 INFLKDFEDKLGITTCSETPLGTAGPLALARDKLDGSGQFFVLNSDVISEYFAE 120

QY 121 LKFKHKGCGEATIMVTKVDESKYGVVMEATGVERFVEKPKIFYGNKINAGIYLLN 180
Db 121 LKFKHKGCGEATIMVTKVDESKYGVVMEATGVERFVEKPKIFYGNKINAGIYLLN 180

QY 181 PSVLDRLELPTSIKEVEPPQIAADQOLYAMVLPGFWMVDVGQPRDYITGLRLYLDLSTKK 240
Db 181 PSVLDRLELPTSIKEVEPPQIAADQOLYAMVLPGFWMVDVGQPRDYITGLRLYLDLSTKK 240

QY 241 SAAKLATGAHVGVNVLVHESAKICEGCLIGDPAIGCCVVEDGVRSLRCTVMRGVRIKK 300
Db 241 SPNRLATGAHVGVNVLVHESAKICEGCLIGDPAIGCCVVEDGVRSLRCTVMRGVRIKK 300

QY 301 HACISNSIIGWSTYGVQWARIENMTILGEDYHVCDEVYSGGVVLPKHEKKSILKPEIV 360
Db 301 HACISNSIIGWSTYGVQWARIENMTILGEDYHVCDEVYSGGVVLPKHEKKSILKPEIV 360

QY 361 M 361
Db 361 M 361

RESULT 2
032287
ID 022287 PRELIMINARY; PRT: 361 AA.
AC 022287;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Putative mannose-1-phosphate guanylttransferase
DE (A02G3970/T517.7)
GN T517.7 OR GMPI OR CYTL.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_Taxid=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Rounsley S.D., Lin X., Ketchum K.A., Crosby M.L., Brandon R.C.,
RA Sykes S.M., Kaul S., Maron T.M., Kerlavage A.R., Adams M.D.,
RA Somerville C.R., Venter J.C.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
[2]
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Weers B., Thornburg R.W.;
RA "Characterization of the cDNA and gene for the Arabidopsis thaliana
RT GDP-mannose pyrophosphorylase";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
[3]
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Lukowitz W., Somerville C.;
RT "Positional cloning of the Arabidopsis cytl gene.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
[4]
RN [4]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Shinn P., Banh J., Bowser L.,
RA Carninci P., Chung M.K., Goldsmith A.D., Hayashizaki Y., Ishida J.,
RA Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B., Lee J.M.,
RA Lin J., Liu S.X., Miranda M., Narusaka M., Nguyen M., Palm C.J.,

```

```

RA Pham P.K., Quach H.L., Sakano H., Sakurai T., Satou M., Seki M.,
RA Southwick A., Toriumi M., Yamada K., Yu G., Shinozaki K., Davis R.W.,
RA Theologis A., Ecker J.R.;
RT "Arabidopsis cDNA clones.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
[5]
RN [5]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis cDNA clones.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
[6]
RN [6]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis cDNA clones.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC003000; AAB87126.1; -
DR EMBL; AF076484; AAC78474.1; -
DR EMBL; AF108060; AAD04627.1; -
DR EMBL; AF361812; AAK32825.1; -
DR EMBL; AF428297; AAL16129.1; -
DR EMBL; AY057541; AAL09781.1; -
DR InterPro; IPR001451; Hexaped_transf.
DR InterPro; IPR001825; NTP_transferase.
DR Pfam; PF00132; hexaped; 3.
DR Pfam; PF00483; NTP_transferase; 1.
DR PROSITE; PS00101; HEXAPEP_TRANSFERASES; UNKNOWN_1.
KW Transferase
SQ SEQUENCE 361 AA; 39577 MW; D282B510E22C2F06 CRC64;
Query Match 90.2%; Score 1681; DB 10; Length 361;
Best Local Similarity 88.6%; Pred. No. 1.3e-129;
Matches 320; Conservative 22; Mismatches 19; Indels 0; Gaps 0;

```

```

QY 1 MKALLVGGFGTGLRPLTSLSPKPLVDFAFKPMILHQIEALKEVGTVTVLAINRPEVM 60
Db 1 MKALLVGGFGTGLRPLTSLSPKPLVDFAFKPMILHQIEALKEVGTVTVLAINRPEVM 60

QY 61 INFLKDFEDKLGITTCSETPLGTAGPLALARDKLDGSGQFFVLNSDVISEYFAE 120
Db 61 INFLKDFEDKLGITTCSETPLGTAGPLALARDKLDGSGQFFVLNSDVISEYFAE 120

QY 121 LKFKHKGCGEATIMVTKVDESKYGVVMEATGVERFVEKPKIFYGNKINAGIYLLN 180
Db 121 LKFKHKGCGEATIMVTKVDESKYGVVMEATGVERFVEKPKIFYGNKINAGIYLLN 180

QY 181 PSVLDRLELPTSIKEVEPPQIAADQOLYAMVLPGFWMVDVGQPRDYITGLRLYLDLSTKK 240
Db 181 PSVLDRLELPTSIKEVEPPQIAADQOLYAMVLPGFWMVDVGQPRDYITGLRLYLDLSTKK 240

QY 241 SAAKLATGAHVGVNVLVHESAKICEGCLIGDPAIGCCVVEDGVRSLRCTVMRGVRIKK 300
Db 241 SPNRLATGAHVGVNVLVHESAKICEGCLIGDPAIGCCVVEDGVRSLRCTVMRGVRIKK 300

QY 301 HACISNSIIGWSTYGVQWARIENMTILGEDYHVCDEVYSGGVVLPKHEKKSILKPEIV 360
Db 301 HACISNSIIGWSTYGVQWARIENMTILGEDYHVCDEVYSGGVVLPKHEKKSILKPEIV 360

QY 361 M 361
Db 361 M 361

```

Db 361 M 361

RESULT 3

Q92TW5 PRELIMINARY; PRT; 361 AA.

AC Q92TW5; 01-MAY-1999 (T-EMBLrel. 10, Created)

DT 01-MAY-1999 (T-EMBLrel. 10, Last sequence update)

DT 01-JUN-2001 (T-EMBLrel. 17, Last annotation update)

DE GDP-mannose pyrophosphorylase.

GN GMP.

OS Solanum tuberosum (Potato).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.

OX NCBI_TaxID=4113;

RN [1]

RP SEQUENCE FROM N.A.

RA Keller R., Kossmann J., Willmitzer L.;

RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF022716; AAD01737.1; -

DR InterPro; IPR001451; Hexapep_transf.

DR InterPro; IPR001825; NTP transferase.

DR Pfam; PF00483; hexapep; 3.

DR Pfam; PF00132; hexapep; 3.

DR PROSITE; PS00101; HEXAPEP_TRANSFERASE; UNKNOWN_1.

DR PROSITE; PS00101; HEXAPEP_TRANSFERASE; UNKNOWN_1.

SQ SEQUENCE 361 AA; 39578 MW; C12A9C4A43BEE4EC CRC64;

Query Match 90.1%; Score 1680; DB 10; Length 361;

Best Local Similarity 87.8%; Pred. No. 1.6e-129;

Matches 317; Conservative 27; Mismatches 17; Indels 0; Gaps 0;

QY 1 MKALILVGGFGTRLRPLTLSPFKPLVDFANKPMILHQIEALKVGVTEVLAINYRPEVM 60

Db 1 MKALILVGGFGTRLRPLTLSPFKPLVDFANKPMILHQIEALKVGVTEVLAINYRPEVM 60

QY 61 INFLLKDFEDKLGITITCSQETPLGTAGPLALARDKLADGSGOPFFVLNSDVISEYPPAE 120

Db 61 INFLLKDFEASLGITITCSQETPLGTAGPLALARDKLADGSGOPFFVLNSDVISEYPPAE 120

QY 121 LIKFHKCHGGEATIMVTKVDEPSKYGVVMEETGRVERFEKPKIFVGNKINAGIYLLN 180

Db 121 LIKFHKCHGGEATIMVTKVDEPSKYGVVMEETGRVERFEKPKIFVGNKINAGIYLLN 180

QY 181 PSVLDRIELRPTSTIEKEVFPQIAADQOLYAMVLPFGFWMVDGQPRDYITGLRLYLDLSIRKK 240

Db 181 PSVLDRIELRPTSTIEKEVFPQIAADQOLYAMVLPFGFWMVDGQPRDYITGLRLYLDLSIRKK 240

QY 241 SSKLATGAHVGNVNLVHESAKIGEGCLIGPDVAIGPGCVVEDGVRLSRTVMRGVRIKK 300

Db 241 SSKLATGAHVGNVNLVHESAKIGEGCLIGPDVAIGPGCVVEDGVRLSRTVMRGVRIKK 300

QY 301 HACISNSIIGWHSTVGQWARIENMTILGEDVHVCDEVYNSGGVLPHPKEIKSSILKPEIV 360

Db 301 HACISNSIIGWHSTVGQWARIENMTILGEDVHVCDEVYNSGGVLPHPKEIKSSILKPEIV 360

QY 361 M 361

Db 361 M 361

RESULT 4

Q9C5B8 PRELIMINARY; PRT; 361 AA.

AC Q9C5B8; 01-JUN-2001 (T-EMBLrel. 17, Created)

DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)

DT 01-JUN-2001 (T-EMBLrel. 17, Last annotation update)

DE GDP-mannose pyrophosphorylase.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RA Avila C.;

RT "Genes responding to phosphate starvation placed together in Arabidopsis genome.";

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ275979; CAC35355.1; -

DR InterPro; IPR001451; Hexapep_transf.

DR InterPro; IPR001825; NTP transferase.

DR Pfam; PF00132; hexapep; 3.

DR Pfam; PF00483; NTP transferase; 1.

DR PROSITE; PS00101; HEXAPEP_TRANSFERASE; UNKNOWN_1.

SQ SEQUENCE 361 AA; 39583 MW; EF36E8FD7D059D9 CRC64;

Query Match 89.8%; Score 1674; DB 10; Length 361;

Best Local Similarity 88.4%; Pred. No. 4.8e-129;

Matches 319; Conservative 22; Mismatches 20; Indels 0; Gaps 0;

QY 1 MKALILVGGFGTRLRPLTLSPFKPLVDFANKPMILHQIEALKVGVTEVLAINYRPEVM 60

Db 1 MKALILVGGFGTRLRPLTLSPFKPLVDFANKPMILHQIEALKVGVTEVLAINYRPEVM 60

QY 61 INFLLKDFEDKLGITITCSQETPLGTAGPLALARDKLADGSGOPFFVLNSDVISEYPPAE 120

Db 61 INFLLKDFEATIMVTKVDEPSKYGVVMEETGRVERFEKPKIFVGNKINAGIYLLN 180

QY 121 LIKFHKCHGGEATIMVTKVDEPSKYGVVMEETGRVERFEKPKIFVGNKINAGIYLLN 180

Db 121 ILEFHKSHGGEASIMVTKVDEPSKYGVVMEETGRVERFEKPKLYVGNKINAGIYLLN 180

QY 181 PSVLDRIELRPTSTIEKEVFPQIAADQOLYAMVLPFGFWMVDGQPRDYITGLRLYLDLSIRKK 240

Db 181 PSVLDRIELRPTSTIEKEVFPQIAADQOLYAMVLPFGFWMVDGQPRDYITGLRLYLDLSIRKK 240

QY 241 SSKLATGAHVGNVNLVHESAKIGEGCLIGPDVAIGPGCVVEDGVRLSRTVMRGVRIKK 300

Db 241 SPKLTSGPHIVGNVNLVDEATIGEGCLIGPDVAIGPGCVESGVRLSRTVMRGVRIKK 300

QY 301 HACISNSIIGWHSTVGQWARIENMTILGEDVHVCDEVYNSGGVLPHPKEIKSSILKPEIV 360

Db 301 HACISNSIIGWHSTVGQWARIENMTILGEDVHVCDEVYNSGGVLPHPKEIKSSILKPEIV 360

QY 361 M 361

Db 361 M 361

RESULT 5

Q94IA7 PRELIMINARY; PRT; 361 AA.

AC Q94IA7; 01-DEC-2001 (T-EMBLrel. 19, Created)

DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)

DT 01-MAR-2002 (T-EMBLrel. 20, Last annotation update)

DE GDP-D-mannose pyrophosphorylase (EC 2.7.7.22).

GN GMPASE.

OS Nicotiana tabacum (Common tobacco).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.

OX NCBI_TaxID=4097;

RN [1]

RP SEQUENCE FROM N.A.

RA Tabata K., Takaoka T., Tateishi Y., Esaka M.;

RT "Gene expression of ascorbic acid-biosynthetic and -oxidizing enzymes and ascorbic acid content in tobacco.";

RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB066279; BAB62108.1; -

DR InterPro; IPR001451; Hexapep_transf.

DR InterPro; IPR001825; NTP transferase.

DR Pfam; PF00132; hexapep; 3.

DR Pfam: PF00483; NTP transferase; 1.
DR PROSITE: PS00101; HEXAPEP_TRANSFERASES; UNKNOWN_1.
KW nucleotidyltransferase; transferase.
SQ SEQUENCE 361 AA; 39572 MW; 40ED456D30B99528 CRC64;

Query Match 89.6%; Score 1671; DB 10; Length 361;
Best Local Similarity 87.3%; Pred. No. 8 5e-129;
Matches 315; Conservative 28; Mismatches 18; Indels 0; Gaps 0;

OY 1 MKALILVGGFGTRLRPLTSLSPKPLVDFAFKPMILHQLAEALKEVGYTEVYLAINYPEYM 60
DB 1 MKALILVGGFGTRLRPLTSLSPKPLVDFAFKPMILHQLAEALKEVGYTEVYLAINYPEYM 60
DB 1 MKALILVGGFGTRLRPLTSLSPKPLVDFAFKPMILHQLAEALKEVGYTEVYLAINYPEYM 60
OY 61 INFLKDFEDKLGITTCSETPEPLGTAGLALARDKLADGSGOPFFVLNSDVISEYPPAE 120
DB 61 INFLKDFEDKLGITTCSETPEPLGTAGLALARDKLADGSGOPFFVLNSDVISEYPPAE 120
OY 121 LKPKHKGCGEATIMVTKVDPSKYGVVMEATGGRVERFVEKPKLFGVGNKINAGIY 180
DB 121 LKPKHKGCGEATIMVTKVDPSKYGVVMEATGGRVERFVEKPKLFGVGNKINAGIY 180
OY 181 PSVLDRIELRPTSLIEKEVFPQIAADQOLYAMVLPFGFMDVGGPRDYITGLRLYLDLSIRKK 240
DB 181 PSVLDRIELRPTSLIEKEVFPQIAADQOLYAMVLPFGFMDVGGPRDYITGLRLYLDLSIRKK 240
OY 241 SAAKLATGAHVGNVNLVHESAKIGEGCLIGPDVAIGPGCVWEDGVRLSRTCVMRGVRIKK 300
DB 241 SAAKLATGAHVGNVNLVHESAKIGEGCLIGPDVAIGPGCVWEDGVRLSRTCVMRGVRIKK 300
OY 301 HACISNSIIGWHSTVGOMARVENMTILGEDVHVCDVGVNCGVVLPHKEIKSSILKPEIV 360
DB 301 HACISNSIIGWHSTVGOMARVENMTILGEDVHVCDVGVNCGVVLPHKEIKSSILKPEIV 360
OY 361 M 361
DB 361 M 361

RESULT 6
OY 09M250 PRELIMINARY; PRT; 364 AA.
AC 09M250;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Mannose-1-phosphate guanylyltransferase-like protein.
GN T22E16.250.

OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Benes V., Wurmbach E., Drzonek H., Ansorge W., Mewes H.W., Lemcke K.,
RA Mayer K.F.X., Quetier F., Salanoubat M.,
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL132975; CAB75917.1; -.
DR InterPro; IPR001451; Hexapep_transf.
DR InterPro; IPR001825; NTP_transferase.
DR Pfam; PF00132; hexapep; 3.
DR Pfam; PF00483; NTP_transferase; 1.
FT PROSITE, PS00101, HEXAPEP_TRANSFERASES; UNKNOWN_1.

KW Nucleotidyltransferase; transferase.
SQ SEQUENCE 364 AA; 39808 MW; 8C2BDAC282CF915D CRC64;
Query Match 85.2%; Score 1597.5; DB 10; Length 364;
Best Local Similarity 83.2%; Pred. No. 6e-122;
Matches 303; Conservative 29; Mismatches 26; Indels 3; Gaps 1;

OY 1 MKALILVGGFGTRLRPLTSLSPKPLVDFAFKPMILHQLAEALKEVGYTEVYLAINYPE-- 58
DB 1 MKALILVGGFGTRLRPLTSLSPKPLVDFAFKPMILHQLAEALKEVGYTEVYLAINYPEQL 60
OY 59 -VMINFLKDFEDKLGITTCSETPEPLGTAGLALARDKLADGSGOPFFVLNSDVISEY 117
DB 61 LVMSKSNVDVEATLIGIKITCSQETPEPLGTAGLALARDKLADGSGOPFFVLNSDVISY 120
OY 119 FAEILKFKHKGCGEATIMVTKVDPSKYGVVMEATGGRVERFVEKPKLFGVGNKINAGIY 177
DB 121 LEEMIAFHNAHGGEASIMVTKVDPSKYGVVMEATGGRVERFVEKPKLFGVGNKINAGIY 180
OY 178 LLNPSVLDRIELRPTSLIEKEVFPQIAADQOLYAMVLPFGFMDVGGPRDYITGLRLYLDLSI 237
DB 181 LLNPSVLDRIELRPTSLIEKEVFPQIAADQOLYAMVLPFGFMDVGGPRDYITGLRLYLDLSI 240
OY 238 RKYSAAKLATGAHVGNVNLVHESAKIGEGCLIGPDVAIGPGCVWEDGVRLSRTCVMRGVR 297
DB 241 RKSPSKLATGPHILGNVLDVDEATIGEGCLIGPDVAIGPGCVWEDGVRLSRTCVMRGVR 300
OY 298 IKKHACISNSIIGWHSTVGOMARVENMTILGEDVHVCDVGVNCGVVLPHKEIKSSILK 357
DB 301 VKRYACISSIIGWHSTVGOMARVENMTILGEDVHVCDVGVNCGVVLPHKEIKSSILK 360
OY 358 EIVM 361
DB 361 DIVM 364

RESULT 7
OY 09M0A3 PRELIMINARY; PRT; 351 AA.
AC 09M0A3;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE GDP-mannose pyrophosphorylase like protein (Fragment).
GN AT4G30570.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Lamar B., Stoneking T., Stumpf J., Mewes H.W., Lemcke K.,
RA Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL161577; CAB79775.1; -.
DR InterPro; IPR001451; Hexapep_transf.
DR InterPro; IPR001825; NTP_transferase.
DR Pfam; PF00132; hexapep; 3.
DR Pfam; PF00483; NTP_transferase; 1.
FT NON_TER 1
SQ SEQUENCE 351 AA; 36681 MW; 86B34E790F5CD3BA CRC64;

Query Match 73.8%; Score 1376; DB 10; Length 351;
Best Local Similarity 76.3%; Pred. No. 1.2e-104;
Matches 264; Conservative 44; Mismatches 36; Indels 2; Gaps 2;

OY 1 MKALILVGGFGTRLRPLTSLSPKPLVDFAFKPMILHQLAEALKEVGYTEVYLAINYPEY 59
DB 1 LKALILVGGFGTRLRPLTSLSPKPLVDFAFKPMILHQLAEALKEVGYTEVYLAINHQQDEV 60
OY 60 MINFLKDFEDKLGITTCSETPEPLGTAGLALARDKLADGSGOPFFVLNSDVISEYPPFA 119
DB 61 MLNFVAEYKELKITFTSETPEPLGTAGLALARDKLADGSGOPFFVLNSDVISEYPELL 120

```
QY 120 ELIKFHCHGGEATIMTKVDEPSKYGVVWMEATGRVERFEKPKIFVGNKINAGIYLL 179
Db 121 EMIEFHTNRRAEASIMTEVDDPSKYGVVWTEEGTARVESFEKPKHEFVGNKINAGIYLL 180
QY 180 NPSVLDRLELRPTSTIEKEVFPQIAADQOOLYAMVLPFGFMDVGQPRDYITGLRLYLDSTRK 239
Db 181 SPSVLDRLELRPTSTIEKEIFPKIASEKKLYAMVLPFGFMDVGQPKDYITGQRMVLSURE 240
QY 240 KSAKLATGAHVGNVLVHESAKIGEGCLIGPDVAIGPGCVVEDGVRSLRSTVMRGVRIK 299
Db 241 KTPDELATGDNIGNVLVHESAVIGEGCLIGPDVVGPGCVIDSGLVFGCTVMRGVWIK 300
QY 300 KHACISNIIHSHSTVGQWARIENNTILGEDVHVCD-EVYSGNGVV 344
Db 301 EHACISNIVGWDSTVGWRARVENITVLGKDVNVADEVYNSGVVI 346

RESULT 8
QY5P6 PRELIMINARY; PRT; 360 AA.
AC QY5P6;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE GDP-mannose pyrophosphorylase B.
GN GMPB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Matthijs G., Schollen E., Dierickx D.;
RT "human homolog of GDP-mannose pyrophosphorylase.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Strausberg R.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RC TISSUE=EYE;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF135421; AAD38516.1; -.
DR EMBL; BC001141; AAH01141.1; -.
DR EMBL; BC008033; AAH08033.1; -.
DR InterPro; IPR001451; Hexapep.transf.
DR InterPro; IPR001825; NTP_transferase.
DR Pfam; PF001132; hexapep; 3.
DR Pfam; PF00483; NTP_transferase; 1.
DR PROSITE; PS00101; HEXAPEP_TRANSFERRASES; UNKNOWN_1.
SQ SEQUENCE 360 AA; 39862 MW; D359AB87459A6B2B CRC64;

Query Match 63.58; Score 1183.5; DB 4; Length 360;
Best Local Similarity 61.24; Pred. No. 7.6e-89;
Matches 221; Conservative 65; Mismatches 74; Indels 1; Gaps 1;

QY 1 MKALILVGGFGTRLRPLTLSPKPLVDFANKPMILHQIEALKVEGVTEVLAINYRPEVM 60
Db 1 MKALILVGGFGTRLRPLTLSTPKPLVDFCNKPLILHQVEALAAAGVDHVLAVTMSQVL 60
QY 61 INFLKDFEDKLGITITCSOETPLGTAGLALARDKLADGSGQPPFVLNSDVISEYFAAE 120
Db 61 EKEMAKQEQRIGIRISHSHEEPLGTAGLALARDLLSE-TADPFVLNSDVICDFPQA 119
QY 121 LIKFHKGGEATIMTKVDEPSKYGVVWMEATGRVERFEKPKIFVGNKINAGIYLLN 180
Db 120 MVQFRRHHGQEGSLVTKVDEPSKYGVVWCEADTGRIHRFEKPKQVFVSNKINAGMITYLS 179
QY 181 PSVLDRLELRPTSTIEKEVFPQIAADQOOLYAMVLPFGFMDVGQPRDYITGLRLYLDSTRK 240
Db 181 PSVLDRLELRPTSTIEKEVFPQIAADQOOLYAMVLPFGFMDVGQPRDYITGLRLYLDSTRK 240
```

```
Db 180 PAVLRRIQLOPTSTIEKEVFPIMAKEGQLYAMELQGFMMDIGQPKDFTGMCLFLQSLRQK 239
QY 241 SAAKLATGAHVGNVLVHESAKIGEGCLIGPDVAIGPGCVVEDGVRSLRSTVMRGVRIK 300
Db 240 QPERLCSGGIVGNVLVDFPSARIGQNCISGPNVSLGPGVVEDGVCIRCTRTVLDRAIRS 299
QY 301 HACISNIIHSHSTVGQWARIENNTILGEDVHVCD-EVYSGNGVVLPHPKESITLKPEIV 360
Db 300 HSWLESCIVGWRRCRVGQWRMENVTVLGEDVIYNDELYLNGASVLPKHSIGESVPEPRII 359
QY 361 M 361
Db 360 M 360

RESULT 9
O74624 PRELIMINARY; PRT; 364 AA.
AC O74624;
DT 01-NOV-1998 (Tremblrel. 08, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE Mannose-1-phosphate guanylttransferase (EC 2.7.13) (MPG1 transferase)
DE (ATP-mannose-1-phosphate guanylttransferase).
GN MPG1.
OS Trichoderma reesei (Hypocrea jecorina).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreales; Hypocreaceae; Hypocrea.
OX NCBI_TaxID=51453;
RN [1]
RP SEQUENCE FROM N.A., AND FUNCTION.
RC STRAIN=ATCC 56765 / QM9414 / RUT C-30;
RX MEDLINE=98309839; Pubmed=9644208;
RA Kruszewska J.S., Saloheimo M., Penttila M., Palamarczyk G.;
RT "Isolation of a Trichoderma reesei cDNA encoding GTP: a-D-mannose-1-
RT phosphate guanylttransferase involved in early steps of protein
RT glycosylation.";
RL Curr. Genet. 33:445-450(1998).
CC -!- FUNCTION: ESSENTIAL PROTEIN WHICH MAY BE INVOLVED IN THE
CC REGULATION OF CELL CYCLE PROGRESSION.
CC -!- CATALYTIC ACTIVITY: GTP + ALPHA-D-MANNOSE 1-PHOSPHATE =
CC DIPHOSPHATE + GDP-MANNOSE.
CC -!- SIMILARITY: TO S.TYPHIMURIUM CDP-GLUCOSE PYROPHOSPHORYLASE (REBF).
DR EMBL; U89991; AAC39498.1; -.
DR InterPro; IPR001451; Hexapep.transf.
DR InterPro; IPR001825; NTP_transferase.
DR Pfam; PF001132; hexapep; 4.
DR Pfam; PF00483; NTP_transferase; 1.
DR PROSITE; PS00101; HEXAPEP_TRANSFERRASES; UNKNOWN_1.
KW Transferase; Nucleotidyltransferase; Kinase; Cell cycle.
SQ SEQUENCE 364 AA; 40284 MW; 66203BAF15284EBA CRC64;

Query Match 63.11%; Score 1176.5; DB 3; Length 364;
Best Local Similarity 61.4%; Pred. No. 2.9e-88;
Matches 224; Conservative 61; Mismatches 75; Indels 5; Gaps 4;

QY 1 MKALILVGGFGTRLRPLTLSPKPLVDFANKPMILHQIEALKVEGVTEVLAINYRPEVM 60
Db 1 MKGLILVGGFGTRLRPLTLTLPLVEFCNKPIMVHQIEALVAAGVTDIVLAVNYRPEIM 60
QY 61 INFLKDFEDKLGITITCSOETPLGTAGLALARDKLADGSGQPPFVLNSDVISEYFAAE 120
Db 61 EKFLAEYEEKYNINIEFSVESEPLDTAGPLKLA-ERILGKDDSPFFVLNSDVICDPFKE 119
QY 121 LIKFHKGGEATIMTKVDEPSKYGVVWME-EATGRVERFEKPKIFVGNKINAGIYLL 179
Db 120 LIEFHKAHGEDEGTIVTKVEEPSKYGVVWHPNHPRIDRFEKPFVEFVGNRNINAGMIIF 179
QY 180 NPSVLDRLELRPTSTIEKEVFPQIAADQOOLYAMVLPFGFMDVGQPRDYITGLRLYLDSTRK 239
Db 180 NPSVLKRIELRPTSTIEKETFPAMVADNQLHSPDLEGFMDVGQPKDFLSGTCLYLSLTK 239
QY 240 KSAAKLA--TGAHV-VGNVLVHESAKIGEGCLIGPDVAIGPGCVVEDGVRSLRSTVMRGV 296
```

```

DB 240 GSKKELTPPTPTPIVGGNNVMIHPSAKIGNKCRIGPNVTIGPDVVYVGGVQLQRCVYLKGS 299
QY 297 RKKKACISNSIIGHSTVGQWARIENMTILGDEHVCDVSNCGVWLPKHKEIKSILK 356
DB 300 KVKDHAHWKSTVGNSTVGRWARLENVVLGDDVTIGDEIYVGGSVLPKHKSIAKANVDV 359
QY 357 PIIVM 361
DB 360 PAIIM 364

RESULT 10
O74404
ID O74484 PRELIMINARY; PRT; 363 AA.
AC O74484; P78779;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DE Probable mannose-1-phosphate guanylyltransferase (EC 2.7.7.13) (ATP-
DE mannose-1-phosphate guanylyltransferase).
GN SPEC1906.01.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Lyne M., Rajandream M.A., Bartoll B.G., Murphy L., Harris D.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=PR745;
RX MEDLINE=98162722; PubMed=9501991;
RA Yoshioka S., Kato K., Nakai K., Okayama H., Nojima H.;
RT "Identification of open reading frames in Schizosaccharomyces pombe
RT CDNA's";
RL DNA Res. 4:363-369(1997).
CC -!- FUNCTION: ESSENTIAL PROTEIN WHICH MAY BE INVOLVED IN THE
CC REGULATION OF CELL CYCLE PROGRESSION (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: GTP + ALPHA-D-MANNOSE 1-PHOSPHATE =
CC DIPHOSPHATE + GDP-MANNOSE.
CC -!- SIMILARITY: TO S.TYPHIMURIUM CDP-GLUCOSE PYROPHOSPHORYLASE (REFB).
DR EMBL: AL031538; CAA20770.1; -.
DR EMBL: D89128; BAAL3790.1; -.
DR InterPro: IPR001451; Hexapep_transf.
DR InterPro: IPR001825; NTP_transferase.
DR Pfam: PF00132; hexapep; 4.
DR PROSITE: PS00101; HEXAPEP_TRANSFERASES; UNKNOWN_1.
KW Transferase; Nucleotidyltransferase; Kinase; Cell cycle.
FT CONFLICT 175 176 GI -> VL (IN REF. 2).
SQ SEQUENCE 363 AA: 39719 NW; 55P55F0ED017C02 CRC64;

Query Match 63.0%; Score 1175; DB 3; Length 363;
Best Local Similarity 62.1%; Pred. NO. 3.8e-88;
Matches 226; Conservative 54; Mismatches 80; Indels 4; Gaps 4;

OY 1 MKALIVGGFGRLRLPTLSFKPKPLVDFANKPMLHIOIALKEVGYTEVVLAINRYPEVM 60
DB 1 MKALIVGGFGRLRLPTLTLPKPLVFENKPKMLHQVLEALAAAGVTDIVLVAINRYPEIM 60
QY 61 INFLKDFEDKLGITITCSQETPLCPAGPLALARDKLADGSGPPFVLNSDVISEYPPFAE 120
DB 61 VEALKKYEKYNVNITFSVENEPLCTAGPLALARDILAK-DHSPFFVLNSDVICEYPPAD 119
QY 121 LKFKHKGGEATIMVTKVDEPSKYGVV-MEATGRTVERFVEKPKIFVGNKKNAGIYLL 179
DB 120 LAAFHKAHCAEGTIVTVKVEEPSKYGVVHYVHNSESLIERFVEKPFVPSNRINGGIVL 179
QY 180 NPSVLDRLELRPTSLEKEVFPQIAADOOILYAMVLPGFWMVDVGOPRDYITGLRLYLSIRK 239

```

```

DB 190 NPSVLDRLELRPTSLEKEVFPQIAADOOILYAMVLPGFWMVDVGOPRDYITGLRLYLSIRK 239
QY 240 KSAAKLA-TGAHVVGNVLVHESAKIGEGCLIGPDVAIGPGCVVDEGVRLSRCTVYMGVRI 298
DB 240 HKEPILAPASSNIIGNVLDIPSATIGNKCKIGPNVIGPNVTIGDVRQLQRCAILKSSRV 299
QY 299 KKHACISNSIIGHSTVGQWARIENMTILGDEHVCDVSNCGVWLPKHKEIKSILK-P- 357
DB 300 RDHAWKSSITVGWNSITLGSWSRLNENSVLGDVVVYVGGSVLPKHKSIAANIEYVP 359
QY 358 EIIVM 361
DB 360 TIIVM 363

RESULT 11
Q9YN61
ID Q9YN61 PRELIMINARY; PRT; 369 AA.
AC Q9YN61;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Cg1123 protein (HL02803P) (R549494P).
GN Cg1123.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballou R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bereman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Chew S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Posler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Iqbal M., Ketchum K.A.,
RA Jalali M., Kalush P., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclab J.M.,
RA Palazotto M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spadling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2195-2195(2000).
RN [2]

```



```
DR InterPro: IPR001825; NTP_transferase.
DR Pfam: PF00132; hexapep; 3.
DR Pfam: PF00483; NTP_transferase; 1.
DR PROSITE: PS00101; HEXAPEP_TRANSFERASES; UNKNOWN_1.
SQ SEQUENCE 387 AA; 42649 MW; CA3BFAD1848F1FE1 CRC64;

Query Match
Best Local Similarity 62.2%; Score 1160; DB 4; Length 387;
Matches 221; Conservative 65; Mismatches 74; Indels 28; Gaps 2;

QY 1 MKALILVGGGTRLRPLTISFPKPLVDFAFKPMILHQIEALKEVGVTEVVLAINRPEVM 60
Db 1 MKALILVGGGTRLRPLTISFPKPLVDFAFKPMILHQIEALKEVGVTEVVLAINRPEVM 60

QY 61 INFLKDPEDKLGITITCSQETEPGLTAGPLALARDKLADGSGOFFVYVLSNDVISEYFPAE 120
Db 61 EXEMKAQEQRLGIRISMHEEPGLTAGPLALARDLSE-TADPFFVLSNDVICDFPQA 119

QY 121 LKFKHCHGGEATIMTKVDPSKYGVVVMEEATGRVERVEPKIPVGNKINAGIYLLN 180
Db 121 LKFKHCHGGEATIMTKVDPSKYGVVVMEEATGRVERVEPKIPVGNKINAGIYLLN 180

QY 120 MVQFRRHGGEGSILVTKVEEPSKYGVVCEADTGRTHREVEKPVQFVSNKINAGMILS 179
Db 120 MVQFRRHGGEGSILVTKVEEPSKYGVVCEADTGRTHREVEKPVQFVSNKINAGMILS 179

QY 181 PSYLDRIELRPTISIEKEVEFQIAADQQLYANVLPFGFMDVQGRDYITGLRLYLDSTRK 240
Db 181 PSYLDRIELRPTISIEKEVEFQIAADQQLYANVLPFGFMDVQGRDYITGLRLYLDSTRK 240

QY 180 PAVLRIRIQLOPSTISIEKEVEFQIAADQQLYANVLPFGFMDVQGRDYITGLRLYLDSTRK 239
Db 180 PAVLRIRIQLOPSTISIEKEVEFQIAADQQLYANVLPFGFMDVQGRDYITGLRLYLDSTRK 239

QY 241 SAAKLATGCAHVGNVNLVHESAKIGEGCLIGPDVAIGPGCVYVEDGVRSLRCTYMRGVRIK 300
Db 241 SAAKLATGCAHVGNVNLVHESAKIGEGCLIGPDVAIGPGCVYVEDGVRSLRCTYMRGVRIK 300

QY 240 QPERLCSGFGIVGNVLYDPFSAIGQNSIGPNVSLGFGVYVEDGVCIRRCTVLRDAIRS 299
Db 240 QPERLCSGFGIVGNVLYDPFSAIGQNSIGPNVSLGFGVYVEDGVCIRRCTVLRDAIRS 299

QY 301 HACISNSIIHSHSTVCOWARIENMTILGEDVHVCDEYVSNNGVVLPHKEIKSSILKPEI 359
Db 301 HACISNSIIHSHSTVCOWARIENMTILGEDVHVCDEYVSNNGVVLPHKEIKSSILKPEI 359

QY 300 HSWLESCIVGWRCVQWVQWVSLWAGLGGEGGECACLPDKAYPLLEVRMENVTVLGEDVIV 359
Db 300 HSWLESCIVGWRCVQWVQWVSLWAGLGGEGGECACLPDKAYPLLEVRMENVTVLGEDVIV 359

QY 334 CDEYVSNNGVVLPHKEIKSSILKPEI 361
Db 334 CDEYVSNNGVVLPHKEIKSSILKPEI 361

QY 360 NDELYLNGASVLPKHSIGESVPEIRIIM 387
Db 360 NDELYLNGASVLPKHSIGESVPEIRIIM 387

RESULT 14
QY725
ID QY725 PRELIMINARY; PRT; 361 AA.
AC QY725;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE GDP-mannose pyrophosphorylase.
GN VIG9.
OS Candida glabrata (Yeast) (Torulopsis glabrata).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5478;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC2001;
RA Ohta A., Sudoh M.;
RT "Candida glabrata VIG9 gene for GDP-mannose pyrophosphorylase.";
DR EMBL: AB020576; BAA77382.1; -
DR InterPro: IPR001451; Hexapep-transf.
DR InterPro: IPR001825; NTP-transferase.
DR Pfam: PF00132; hexapep; 4.
DR Pfam: PF00483; NTP-transferase; 1.
DR PROSITE: PS00101; HEXAPEP_TRANSFERASES; UNKNOWN_2.
SQ SEQUENCE 361 AA; 39298 MW; E255084D2EDB2A1D CRC64;

Query Match
Best Local Similarity 60.1%; Score 1121; DB 3; Length 361;
Matches 211; Conservative 62; Mismatches 87; Indels 2; Gaps 2;

QY 1 MKALILVGGGTRLRPLTISFPKPLVDFAFKPMILHQIEALKEVGVTEVVLAINRPEVM 60
Db 1 MKALILVGGGTRLRPLTISFPKPLVDFAFKPMILHQIEALKEVGVTEVVLAINRPEVM 60
```

Query Match

Score 1099.5; DB 3; Length 362;

```
QY 61 INFLKDPEDKLGITITCSQETEPGLTAGPLALARDKLADGSGOFFVYVLSNDVISEYFPAE 120
Db 61 VETLQKYKEKGVSTISFVETEPGLTAGPLKLA-EKVLKDKNSPFVLSNDVICYPFKE 119

QY 121 LKFKHCHGGEATIMTKVDPSKYGVVVMEEAT-GRVERVEKPKIFVGNKINAGIYLL 179
Db 121 LADPHKAGGAGTIVATKVDPSKYGVVHDIATFNLDIRFVEKPKFVGNRINAGLXIL 179

QY 180 NPSYLDRIELRPTISIEKEVEFQIAADQQLYANVLPFGFMDVQGRDYITGLRLYLDSTRK 239
Db 180 NPEVDLTIEKPTSTETETFDILVEOKSLYSFOLGFWMDVGPKDFLSGTVLVLNSVSK 239

QY 240 KSAKLATGCAHVGNVNLVHESAKIGEGCLIGPDVAIGPGCVYVEDGVRSLRCTYMRGVRIK 299
Db 240 KNPEKLTAKGDNIVGNVMDPSAKIAASAKVGPDDVVGIPNVITGCVRTSRVLSDSSTQ 299

QY 300 KHACISNSIIHSHSTVCOWARIENMTILGEDVHVCDEYVSNNGVVLPHKEIKSSILKPEI 359
Db 300 DHSLVKSTIVGKSTVGKWCRLGTVLGDVYVNGKYLPHKHSISANVPEBAI 359

QY 360 VM 361
Db 360 IM 361

RESULT 15
QY3827
ID QY3827 PRELIMINARY; PRT; 362 AA.
AC QY3827;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Mannose-1-phosphate guanylyltransferase (EC 2.7.7.13) (ATP-mannose-1-phosphate guanylyltransferase) (CASRbl).
GN VIG9 OR SRB1 OR PSAL.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1161;
RX MEDLINE=38455809; Pubmed=3782409;
RA Marit S., Walmsley R.M., Staveva L.I.;
RT "Cloning and sequencing of the Candida albicans homologue of SRB1/PSAL/VIG9, the essential gene encoding GDP-mannose pyrophosphorylase in Saccharomycetes cerevisiae.";
RL Microbiology 144:2417-2426(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=IFO1060;
RA Ohta A., Sudoh M.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: ESSENTIAL PROTEIN WHICH MAY BE INVOLVED IN THE REGULATION OF CELL CYCLE PROGRESSION.
CC -1- CATALYTIC ACTIVITY: GTP + ALPHA-D-MANNOSE 1-PHOSPHATE -> DIPHOSPHATE + GDP-MANNOSE.
CC -1- SIMILARITY: TO S.TYPHIMURIUM CDP-GLUCOSE PYROPHOSPHORYLASE (RFBF).
DR EMBL: AF030299; AAC64911.1; -
DR EMBL: AF030300; AAC64912.1; -
DR EMBL: AB020596; BAA34807.1; -
DR InterPro: IPR001451; Hexapep-transf.
DR InterPro: IPR001825; NTP-transferase.
DR Pfam: PF00132; hexapep; 4.
DR Pfam: PF00483; NTP-transferase; 1.
KW Transferase; Nucleotidyltransferase; Kinase; Cell cycle.
FT VARIANT 197 198 ET -> DP (IN STRAIN 1161).
FT CONFLICT 289 289 Q -> R (IN AAC64912).
FT VARIANT 333 333 E -> Q (IN STRAIN 1161).
SQ SEQUENCE 362 AA; 39975 MW; 052DCC32D0E16923 CRC64;

Query Match
Score 59.08; Length 362;
```


Best Local Similarity 57.9%; Pred. No. 5.9e-82;			
Matches	210; Conservative	60; Mismatches	90; Indels
3; Gaps			
QY	1	MKALILVGGFGTRLRLPLTLSPKPLVDFANKPMTLHQIEALKEVGVTEVWLAINYRPEVM	60
Db	1	MKGLILVGGYGTRLRLPLTLPLKPLVEFGNRPMTLHQIEALAAAGVTDIVLAVNIRPEVM	60
QY	61	INFLKDPEDKLGITITCSQETPLGTAGLALARDKLADSGSQPFFVLNSDVISEYPPAE	120
Db	61	VSTLKKYEEYGVISITFSEVEEPLGTAGPLKLAEEVLKK-DDSPFFVLNSDVICDYPKE	119
QY	121	LIRFKHCHGGEATIMVTKVDEPSKYGVVWMEAT-GRVEREVEKPKIFVGKNKINAGIYLL	179
Db	120	LADFHKAHGAAGTIVATKVDEPSKYGVIVHDDTPNLIDREVEKPVVEVGNRINAGLYIL	179
QY	180	NPSVLDRIELRPTSIEKEVPPQIAADQOLYAMVLPGFWMVDYQPRDYITGLRLYLDLSRK	239
Db	180	NPSVIDLIEMRPTSIEKETFPILVEQKOLYSFDLEGYWMVDYQPRDFUSGTCCLYLTLSLK	239
QY	240	KSAAKLATGAHV-VGNVLVHESAKIGEGCLIGPDVAIGPGCWVEDGVRLSCTVMRGVRI	298
Db	240	KHPEKLCCKEYVHGGNVLIIDPTAKIHPKSALIGPNVTIGPNVVGEGARIQSVLLANSQV	299
QY	299	KKHACISNSTIGWHSTYQWARIENMTILGEDVHVCDEVYSGNGVVLPHKSIKPE	358
Db	300	KDHAWKSTIVGMNSRIGKWARTEGTVLGGDVEVKNEIYVNGAKVLPKHSISSNVEKES	359
QY	359	IVM	361
Db	360	IIM	362

Search completed: November 27, 2002, 05:09:35
Job time : 85 secs